NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

AN ANALYSIS OF HELICOPTER PILOT SCAN TECHNIQUES WHILE FLYING AT LOW ALTITUDES AND HIGH SPEED

by

Christopher E. Kirby

September 2012

Thesis Advisor:	Quinn Kennedy
Thesis Co-Advisors:	Ji Hyun Yang
	Joseph Sullivan

Approved for public release, distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2012	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE An Analysis of Helicopter Pilot Scan Techniques While Flying at Low Altitudes and High Speed			**5. FUNDING NUMBERS**
6. AUTHOR(S) Christopher E. Kirby			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number NPS20120011.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

This study compared how non-experienced and experienced pilots reacted in terms of their scan patterns during a simulated high speed low level flight. The focus of this study was specifically on the flight regimes encountered by helicopter pilots. Information obtained from this research may aid training effectiveness specific to helicopter aviation.
Methods: There were 17 military officers, all active-duty Navy helicopter pilots, who all had different levels of flight experience based on their total flight times. Each pilot was asked to successfully fly and navigate a course through a simulated southern Californian desert in a fixed-based helicopter simulator modeled after the U.S. Navy's MH-60S. The location of their scan was tracked by an eye-tracking system in order to determine scan rate and locations while they flew the course. All of the flight parameters, such as airspeed and altitude, were recorded by the simulator's recording system.
Results: Analysis of the results obtained from the eye tracking system indicated a decreasing relationship between scan rate and pilot experience, indicating that the scan rate decreases as a pilot becomes more experienced. The analysis uses altitude variance as a measure of performance. Results indicate that higher scan rates correlate with higher degrees of variance in the altitude, indicating that a quicker scan does not necessarily result in better performance. The higher experienced pilots show a lower altitude variance overall (they were more consistent in maintaining a constant altitude above the ground), yet those pilots all exhibited slower scan rates.
Discussion: The integration of the eye tracking technology with a simulator representing an aircraft currently in service was a success. Although none of the null hypotheses presented were rejected, trends were evident in scan rates when compared with pilot experience. The relatively small sample size was identified as the major causal factor for the lack of significance.

14. SUBJECT TERMS Scan, Expertise, Simulator, Low Level Flight, Helicopter			**15. NUMBER OF PAGES** 119
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution unlimited

AN ANALYSIS OF HELICOPTER PILOT SCAN TECHNIQUES WHILE FLYING AT LOW ALTITUDES AND HIGH SPEED

Christopher E. Kirby
Commander, United States Navy
B.A., State University of New York at Albany, 1996

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2012

Author: Christopher E. Kirby

Approved by: Quinn Kennedy
Thesis Advisor

Ji Hyun Yang
Thesis Co-Advisor

Joseph Sullivan
Thesis Co-Advisor

Kip Smith
Second Reader

Robert F. Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This study compared how non-experienced and experienced pilots reacted in terms of their scan patterns during a simulated high speed low level flight. The focus of this study was specifically on the flight regimes encountered by helicopter pilots. Information obtained from this research may aid training effectiveness specific to helicopter aviation.

Methods: There were 17 military officers, all active-duty Navy helicopter pilots, who all had different levels of flight experience based on their total flight times. Each pilot was asked to successfully fly and navigate a course through a simulated southern Californian desert in a fixed-based helicopter simulator modeled after the U.S. Navy's MH-60S. The location of their scan was tracked by an eye-tracking system in order to determine scan rate and locations while they flew the course. All of the flight parameters, such as airspeed and altitude, were recorded by the simulator's recording system.

Results: Analysis of the results obtained from the eye tracking system indicated a decreasing relationship between scan rate and pilot experience, indicating that the scan rate decreases as a pilot becomes more experienced. The analysis uses altitude variance as a measure of performance. Results indicate that higher scan rates correlate with higher degrees of variance in the altitude, indicating that a quicker scan does not necessarily result in better performance. The higher experienced pilots show a lower altitude variance overall (they were more consistent in maintaining a constant altitude above the ground), yet those pilots all exhibited slower scan rates.

Discussion: The integration of the eye tracking technology with a simulator representing an aircraft currently in service was a success. Although none of the null hypotheses presented were rejected, trends were evident in scan rates when compared with pilot experience. The relatively small sample size was identified as the major causal factor for the lack of significance.

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AD	Aircraft Diagnostics
AGL	Above Ground Level
AOI	Area of Interest
CDTI	Cockpit Display of Traffic Information
CFIT	Controlled Flight Into Terrain
CSV	Comma Separated File
EMS	Emergency Medical Service
FAA	Federal Aviation Administration
GPS	Global Positioning Satellites
GS	Ground Speed
HEMS	Helicopter Emergency Medical Service
HS	Helicopter Anti-Submarine
HSC	Helicopter Sea Combat
HUD	Heads-Up Display
IAS	Indicated Air Speed
IFR	Instrument Flight Rules
IM	Instrument Display to Map
IMC	Instrument Meteorological Conditions
IO	Instrument Display to Out-of-the-Window
IRB	Internal Review Board
KIAS	Knots Indicated Air Speed
MFD	Multi-Function Display
MI	Map to Instrument Display
MO	Map to Out-of-the-Window
NAS	Naval Air Station
NASNI	Naval Air Station North Island
NPS	Naval Post-Graduate School
OI	Out-of-the-Window to Instrument Display
OM	Out-of-the-Window to Map
OTW	Out-of-the-Window

PI	Principal Investigator
RA	Research Assistant
RMS	Root Mean Squared
SDTM	Simulation and Training Device Manager
SV	Synthetic Vision
TFH	Total Flight Hours
TOFT	Tactical Operational Flight Trainer
VFR	Visual Flight Rules
VMC	Visual Meteorological Conditions

EXECUTIVE SUMMARY

Helicopters operate at low altitude levels (within 500 feet of ground level); it is the nature of their mission and the way they survive in combat environments. As helicopters evolve and become faster and more agile, their pilots will be expected to navigate at low altitude levels while traveling at high speeds. The primary danger in these missions is Controlled Flight Into Terrain (CFIT).

The ability of a pilot to interpret information from a combination of sources determines the success of a mission, as well as the survival of the aircraft and its crew. These sources include the outside environment (the visual scan), the instrument panel (flight profile information), displays that inform the pilot of the aircraft's status, and additional information from navigation charts or Global Positioning Satellites (GPS) displays. A competent pilot is able to move his or her scan from source to source in such a way that maximizes the assimilation of information, and react accordingly to safely maneuver the aircraft.

This thesis investigates the relationship between a helicopter pilot's experience level and his or her visual scan patterns during high-speed low-level flight. The research focuses on scanning patterns in flight regimes that are not optimal. For helicopters, flying at high speeds and low-levels is not the safest way to fly, but in times of war, it is necessary for survival. A helicopter's primary means of defense while flying in combat is to remain low and masked by the terrain. Keeping a high rate of speed is vital for reducing the time an enemy has to target the helicopter as it passes over. The combined low-level, high-speed flight results from these two needs.

Seventeen military officers, all active-duty Navy helicopter pilots, who all had different levels of flight experience based on their total flight times, volunteered for the study. Each pilot was asked to successfully fly and navigate a course through a simulated southern Californian desert in a fixed-based helicopter simulator modeled after the U.S. Navy's MH-60S. The location of their scan was tracked by an eye-tracking system in

order to determine scan rate and locations while they flew the course. All of the flight parameters, such as airspeed and altitude, were recorded by the simulator's recording system.

Analysis of the results obtained from the eye tracking system indicates a decreasing relationship between scan rate and pilot experience, indicating that the scan rate decreases as a pilot becomes more experienced. The analysis uses altitude variance as a measure of performance. The results indicate that higher scan rates correlate with higher degrees of variance in the altitude, indicating that a quicker scan does not necessarily result in better performance. Pilot fixation events–those events in which a pilot looked at an area of interest for more than 70 milliseconds–were also analyzed. Exploratory analyses revealed that the amount of "no fixation" events significantly decrease with regard to pilot experience. The higher experienced pilots show a lower altitude variance overall (they were more consistent in maintaining a constant altitude above the ground), yet those pilots all exhibited slower scan rates.

The research focused on scanning patterns in flight regimes that are not optimal: flight at high speeds and low altitude levels. Our primary goal was to gain an understanding of the unique scan characteristics that might present themselves in this challenging arena of flight by investigating the relationship between a helicopter pilot's experience level and his or her visual scan patterns during high-speed, low-level flight. Although none of the null hypotheses presented were rejected, trends were evident in scan rates when compared with pilot experience. The relatively small sample size was identified as the major causal factor for the lack of significance. Data was lost on a total of five subjects from both the eye tracking system and the simulator's flight recording software.

A secondary goal of this thesis was to verify that FaceLab can be adapted for use in a simulator that is not in the laboratory environment. The data from FaceLab and the simulator was combined to complete the analysis outlined in this study. The integration of the eye tracking technology with a simulator representing an aircraft currently in service was a success.

ACKNOWLEDGMENTS

This thesis involved two departments at the Naval Postgraduate School, a helicopter wing, three squadrons, a contractor and numerous other people who contributed one way or another. The outpouring of support from all involved has consistently motivated me to produce the best product I could, with the hopes of creating a product that contributes to aviation safety.

First and foremost, there is my family. My wife, who inspired me to attempt a master's degree in the first place, has been a strong support and inspiration throughout this process. From coordinating trips to San Diego to taking one last look at the finished product, her support has defined what love is. My son's non-stop inquisitiveness and adventurous spirit moved me to keep asking questions and try to see things through his eyes. Both my wife and my son have been most patient with me as I struggled through school and this project.

The staff at the Naval Postgraduate School went above and beyond what I considered the necessary support for a thesis. My deepest thanks go to Professors Quinn Kennedy and Ji-Hyun Yang for allowing me to run down to San Diego with their precious equipment to try out this crazy idea. Their constant guidance was necessary and most helpful to a helicopter pilot trying to swim through the academic channel. CDR Joseph Sullivan has the same passion for improving aviation and provided mentorship in that light throughout. Dr. Kip Smith was our human factors expert. He kept the "human" in this project by keeping the focus on the human-machine interaction that occurs in every cockpit. None of the travel would have occurred without the help of Mr. Joe Hyde in the NPS Travel Office. Finally, to the very helpful research assistant, Mr. Jesse Huston, who kept the whole experiment from failing in San Diego through his genius use of duct tape and construction paper.

The outpouring of support from the wing and the squadrons at NAS North Island was staggering. The Commodore of the Helicopter Sea Combat Wing Pacific was our largest benefactor when we were operating on the air base. The Commanding Officers of

Helicopter Sea Combat Squadron Three, Helicopter Anti-Submarine Squadron Four and Helicopter Sea Combat Squadron Two-One all showed their support by allowing their pilots to take time from their busy operational and training schedules to participate in this study. Mr. Carl Jacobs, the wing's simulation and training device manager, and his staff offered overwhelming technical support concerning TOFT-2. A large "thank you" to CDR Jeff "Angry" Nolan, Chief Staff for the Helicopter Sea Combat Wing Pacific for tolerating my endless visits to his office seeking advice on how to make the whole deal work.

I thank God for all of these people, for those I have not mentioned and for the opportunities He gave and continues to give me.

Proverbs 14:23–All hard work brings a profit, but mere talk leads only to poverty.

I. INTRODUCTION

Helicopters operate at low altitude levels (within 500 feet of ground level); it is the nature of their mission and the way they survive in combat environments. As helicopters evolve and become faster and more agile, their pilots will be expected to navigate at low altitude levels while traveling at high speeds. The primary danger in these missions is Controlled Flight Into Terrain (CFIT).

The ability of a pilot to interpret information from a combination of sources determines the success of a mission, as well as the survival of the aircraft and its crew. These sources include the outside environment (the visual scan), the instrument panel (flight profile information), displays that inform the pilot of the aircraft's status, and additional information from navigation charts or Global Positioning Satellites (GPS) displays. A competent pilot is able to move his or her scan from source to source in such a way that maximizes the assimilation of information, and react accordingly to safely maneuver the aircraft.

This thesis investigates the relationship between a helicopter pilot's experience level and his or her visual scan patterns during high-speed, low-level flight. The research focuses on scanning patterns in flight regimes that are not optimal. For helicopters, flying at high speeds and low levels is not the safest way to fly, but in times of war, it is necessary for survival. A helicopter's primary means of defense while flying in combat is to remain low and masked by the terrain. Keeping a high rate of speed is vital for reducing the time an enemy has to target the helicopter as it passes over. The combined low-level, high-speed flight results from these two needs.

In the case of an emergency situation in which the pilot is alerted to a condition that threatens the aircraft's ability to safely complete the flight, the information from the display describing the nature of the aircraft emergency needs to be scanned and processed on top of all of the normal flight operating parameters. The pilot is still expected to fly, navigate and communicate throughout the situation in which the aircraft is not operating

correctly. While operating in this flight regime, information needs to be assimilated quickly and efficiently to account for the decreased reaction time demanded in this environment.

As a pilot increases in experience, his or her scan should become more efficient at gathering the needed information to safely fly and navigate. This thesis seeks to quantify what constitutes a more efficient scan.

A. BACKGROUND

The year 2008 proved to be the deadliest year on record for Emergency Medical Service (EMS) helicopter aviation. Six months into the year, 16 people had perished in EMS helicopter crashes. The accident rate for the year for helicopters was 6000 times greater than fixed-wing commercial air carriers (Evans, 2008).

Bitton (2008) highlighted the need to further understand the scan patterns of helicopter pilots by describing the ongoing research efforts by the FAA and the EMS helicopter industry to study and analyze helicopter emergency medical service (HEMS) accidents, in order to better understand the causes (Bitton, 2008): “The primary cause of accidents remains Controlled Flight Into Terrain (CFIT) and inadvertent flight into clouds or diminishing visibility, with the results of loss of situational awareness.”

B. LITERATURE REVIEW

Duquette and Dorr (2010) were also conducting the same kind of research in conjunction with the FAA in order to better understand the alarmingly high accident trend in 2008. The results of this ongoing investigation are that the main causes of HEMS accidents were controlled flight into terrain (CFIT), inadvertent operation into instrument meteorological conditions (IMC), and pilot spatial disorientation/lack of situational awareness in night operations. The Honeywell Corporation conducted a CFIT study in order to dispel some common helicopter accident myths. Importantly, they found that the majority of CFIT (52%) accidents occur during the daylight hours in visual meteorological conditions (VMC) (61%) (Learmount, 2005). The study also found that 67% of helicopter CFIT incidents occur during the cruise phase of flight.

Few studies had been done to date on helicopter pilots in these regimes. Given the alarming accident statistics above, a study of the scan patterns of helicopter pilots is long overdue (Crognale, 2008). Recently, the use of synthetic vision (SV) and a heads-up-display (HUD) have been a topic of discussion in the aviation community. Synthetic vision uses external cameras to provide the pilot with an enhanced view of the outside world, usually with the assistance of night vision technology. In conjunction with the use of a HUD, or a display mounted in such a way that the pilot can get information without removing his or her gaze from the windscreen, synthetic vision could greatly increase aviation safety by aiding helicopter pilots in scanning for hazards during flight. A HUD would be designed in such a way that presents important and time-sensitive information to the pilot in one place, without pilots having to shift their scan from the outside environment. This set up allows pilots to remain focused on an important task, such as landing, while still receiving information that contributes to situational awareness.

Wickens (2001) states that the effort required to shift a gaze or direct the attention over to another area that may be a long distance from the area currently scanned can sometimes inhibit that shift or re-direction. Wickens (2001) further states: “We have found, in a rotorcraft simulation, that directing the pilot’s attention to important hazards like power lines or terrain by cueing, while offering benefits to the detection of those hazards, will direct attention away from other un-cued hazards in the area.” A HUD, acting as a centralized location for all the required information while still allowing the pilot access to the outside environment, could eliminate that inhibition.

In *Rotor and Wing*, Adams (2010) argues for the use of “overlapping” displays, such as multi-function displays (MFD), that reduce the need for pilots to move their attention from one area to another. He advocates for the integration of synthetic vision (SV) and enhanced vision (EV) with helicopter system and navigation information. Much of this technology is already available on the fixed wing market, and is heavily advocated by the FAA through operational credits to those carriers employing these technologies. The FAA intends to extend these operational credits so that they include rotary-winged aircraft (Adams, 2010). The FAA also requires that fixed wing commercial air carriers

have two independent traffic, terrain and obstacle warning systems installed in their aircraft. This requirement does not extend to rotary-wing commercial air carriers such as EMS helicopters (Evans, 2008).

The need for better displays in helicopters has been established, but in order to understand what helicopter pilots need on those displays, their rate of scan and scanning pattern have to be understood. In order to properly design the displays in the spirit of the machine supporting the operator, experiments recording helicopter pilots' scan patterns during all phases of flight are needed. The purpose of this study was to begin to fill in this gap in knowledge.

1. Previous Studies Investigating Eye Tracking and Aviator Flight Performance

Several studies have investigated visual scan patterns in aviator flight performance. In this section, I focus on the ones most relevant to the current study. Mumaw, Sarter, and Wickens (2001) studied 20 Boeing 747-400 pilots in a simulator in order to understand the role of "pilot monitoring in the loss of awareness on automated flight decks." Prior studies indicated unsafe flying conditions when pilots fail to understand or have some confusion over an aircraft's automated flight systems. Mumaw et al. (2001) attempted to gain an understanding of the interaction between pilots and an aircraft's automated flight systems through the use of an eye-tracking system. The initial analysis of the eye-tracking data focused on how fixations were distributed in each area of interest (AOI). From this information, the investigators were able to determine where the pilots were looking in different phases of the flight. Different traffic schemes and emergency scenarios were also introduced. Mumaw et al. (2001) found that many pilots had ineffective scan patterns, and the type of automation feedback used in the aircraft had caused unnecessary fixations on areas of little importance. Given these results, Mumaw et al. (2001) were able to suggest improvements in pilot training until new automation and instrumentation interfaces could be developed.

Two papers by Wickens et Goh, Helleberg and Talleur. (2002, 2003) center around their study in which a sample of 12 pilots flew a full mission simulator to examine

two aspects of advanced aviation display technology, the digital data-link and the cockpit display of traffic information (CDTI). The study used a high-fidelity flight simulator based on a Frasca 142. Pilots wore a head-mounted eye-tracking system throughout the flight. The information regarding other aircraft in the area was presented to the pilots through auditory cues, visual cues, and a combination of both.

They found, through the use of eye-tracking technology, that pilots used a variety of outside scan patterns, with the “sector sweep” method being the most proficient. Pilots were able to detect traffic more efficiently without disrupting piloting tasks using the sector sweep method. Percentage dwell time (PDT) time on areas of interest (OW, IP, and CDTI (Cockpit Digital Traffic Indicator)) was calculated via mean dwell duration (MDD) using the eye trackers to show how much time each pilot spent on a particular area of interest. The study was able to show that the auditory traffic warning resulted in the pilots’ spending more time looking out of the window (OW) for traffic, rather than trying to interpret what a visual display was trying to tell them (Wickens et al., 2002). The results showed different types of scan patterns depending on the traffic load.

Whereas Mumaw et al. (2001) and Wickens et al. (2002, 2003) examined visual scan patterns among a sample of highly skilled pilots, most other studies investigating visual scan patterns and flight performance have focused on expertise differences. For example, Bellenkes, Wickens and Kramer (1997) measured attention control by analyzing visual scanning behavior during a simulated VFR flight in expert and novice pilots. A novice pilot was defined as having logged between 40 and 70 hours of VFR flight time. The expert pilots were rated Air Force pilots having logged between 1500 hours and 2150 hours of flight time. It was a ground-breaking study in which attention control was broken into perception (which channels to select) and response (which actions to perform). Visual scanning was also used in an attempt to describe pilots' mental models. Three challenges to attention were found. The first is that attention is limited, so pilots have to prioritize it. Secondly, when the main axes were “sluggish,” proactive attention was required well in advance of changes. Lastly, complex interactive dynamics require increased attention. The most important instrument, in terms of the amount of time the pilots spent looking at it, was the attitude indicator. Results indicate

clear expertise differences in visual scan patterns. Experts scanned the flight instruments more often than novices, and novices had longer dwell times than experts. In particular, experts scanned the directional gyro and altimeter more often than novices. Novices dwelled longer on the vertical speed indicator and the turn coordinator than experts. Experts performed better than novices in terms of altitude control, particularly on the two most difficult segments of the route. They also had fewer airspeed errors overall.

Another study by Karsarkis,Stehwien, Hickox, Aretz and Wickens (2001) posed the following research question: To what extent are various aviation stimuli visually sampled, and how do scanning strategies differ between novice and expert pilots? Novice pilots were ten U.S. Air Force Academy cadets who had 40–70 hours of VFR flight time. Experts were six Air Force pilots with 1500–2150 flight hours. This study was similar to one conducted by Bellenkes et al. (1997) in which they found different scan patterns for turning, climbing, and descending. Karsarkis et al. presented the hypothesis that expert pilots would have shorter dwell times and more fixations on all instruments, particularly on the airspeed indicator, altimeter, and vertical velocity indicator, and will look out the window more often than novices. They also hypothesized that landing performance correlated with total fixations and dwell time per fixation. From the data analysis in the Karsarkis study, poor landings were defined as those that were below the median of all landings across all subjects. Pilot fixations were classified into four visual areas of interest: the upper 2/3 of computer screen, and as the trial progressed this included the runway; the airspeed indicator, the altimeter, and other instruments on the far right side of instrument panel. The results were as expected. Pilots spent much more time looking out of the window, indicating (at least in part) why experts performed better landings: experts paid more attention to airspeed than altimeter, with more fixations outside. These outside fixations tended to be at a more distant point on the runway; experts also had shorter dwells (and more fixations) on almost all instruments and on runway. The authors suggest that the focus on airspeed is a key strategy, and this strategy was particularly evident when altitude was changing. The other part was that experts had shorter dwells on everything, indicating automation and that the experts have more time to scan other locations. Importantly, more fixations and shorter dwell times

also were associated with good landings, providing indirect evidence that these visual strategies cause expertise differences in landing performance.

The studies by Bellenkes et al. and Karsarkis et al. are good examples of the evolution of eye-tracking technology and data analysis, as well as how eye-tracking can be used to detect successful visual scan strategies used by expert pilots. Bellenkes et al. assessed eye patterns during IFR cruise flight, while Karsarkis et al. focused primarily on flight under visual flight conditions. A study by Ottati Hickox and Richter (1999), which was based on work by Bellenkes et al. (1997), tested the hypothesis that expert pilots spend less time finding and fixating on individual landmarks and are able to use landmarks to navigate more accurately than novice pilots. They also tested the hypothesis that pilots classified as novices will have a harder time finding landmarks and thus have longer dwell times than pilots classified as experts. The experiments relied on 20 cadets from the U.S. Air Force Academy. Half of the cadets were considered experienced pilots; experienced pilots were those having a private pilot's license and at least 50 flight hours. Novice pilots had between 5 and 15 flight hours logged. The participants had simulated flights over a 40 nautical mile route that consisted of five distinct checkpoints. The pilots could utilize a 90-degree forward field of view map representation of their current location. Eye tracking data consisting of those fixations that lasted longer than 0.1 seconds were analyzed. Fixation dwells were defined as those that lasted at least one second. The study found that novice pilots were more likely to fly out of the window, rather than relying on instrumentation to guide them through the 40 nautical mile route. Experienced pilots had more fixations than novices, but no differences in dwell time. This study suggests that the fixations of the expert pilots were more deliberate. The authors conclude that novice pilots are more likely to use spontaneous fixations during flight tasks in order to gain an accurate orientation.

Only one study was found, other than the ones conducted here at NPS described below (Sullivan, Yang, Day and Kennedy, 2011), that used eye-tracking technology for low-level en-route flight in a helicopter. The first project found that examined the visual workload of the navigator/copilot during terrain flight in a UH-1H helicopter was conducted in 1979 by Sanders, Simmons and Hoffman. Visual performance was

measured using external eye-tracking measures (called an occulometer). The study did not focus on the flying pilot, but rather on the navigator. Sanders et al. found that helicopter navigators spent different amounts of time looking out of different view screens to gather data for navigation. Navigators spent 46.8% of the time looking out the left windscreen, 5% through the right view screen, 4.9% of the time looking out the left gunner's window, while the rest of the time was spent looking at the hand held map used to navigate. Sanders' study laid the groundwork for the type of analysis done in the results section of this thesis, which involved analyzing the amount of time a pilot looked at a particular area of interest and relating the scan pattern to a performance measure.

2. Studies Utilizing Altitude as a Measure of Pilot Performance

Helicopter overland navigation while flying at low altitude levels is a demanding task for pilots as it entails additional tasks that do not involve the simple control of the aircraft. Sullivan et al. (2001) found that a common flight performance measure, RMS error of flight trajectory, does not predict expertise levels in helicopter overland navigation as it does in other aviation tasks. They found that helicopter pilots are trained to adapt their en-route or "between way points" navigation solution based on what they are seeing in terms of terrain at the current time. Thus, in evaluating helicopter pilots, a different measure of expertise beyond RMS error was needed. This thesis uses the amount of deviation from the assigned altitude parameters to evaluate pilot performance.

The concept of using altitude deviation as a measure of performance is not a ground-breaking and new practice. A study done in 1965 by Soliday and Scohan used altitude deviation as a measure of performance in order to determine the effects of task loading on pilots. Only three pilots were used, and were asked to fly at 500 feet above the ground in a light fighter aircraft. The results were as expected; as the terrain became more difficult, or as the pilots were asked to fly at higher airspeeds, the altitude deviations increased in both severity and number. It was also discovered that pilots fly at different altitudes depending on how they approach a slope of a mountain (higher going up, lower coming down).

Wickens et al. (2003) also used altitude deviation as a measure of performance. They found a positive association between altitude error and traffic load. Finally, Yesavage, Otto Leirer, Denari and Hollister (1985) conducted a study to examine THC carry-over effects on a simple piloting task 24 hours after smoking of the drug. Pilots (both those who smoke and those who did not) were asked to fly a simulator that mimicked the controls and cockpit of a Cessna 172. Different parameters were measured to determine pilot performance, to include "average lateral deviation from an ideal glideslope," or altitude deviation. Through this measure of performance, it was discovered that the effects of THC were most notable one hour after smoking the drug. Pilots who smoked THC deviated significantly more than at baseline. After 24 hours from smoking, the effects subsided (the average altitude deviation did not differ significantly from the baseline).

3. Literature Summary

The need for research involving the scanning habits of helicopter pilots exists. Many of the studies, some of which are listed here, focus primarily on the scanning patterns of fixed wing pilots. Further research is needed to fully understand the scanning patterns of helicopter pilots, whose operating conditions are far different than those of fixed wing pilots. Scan patterns change as a pilot's experience changes. Does this carry over to helicopter pilots? If there are performance differences, how do the helicopter pilots that exhibit a better performance scan?

The studies outlined above use eye-tracking technology in one way or another to fulfill their research goals. The goal of the study conducted for this thesis is to examine scan techniques used by helicopter pilots across a wide range of experience levels. This thesis differs from the previous research in that it focuses on pilots flying helicopters in cruise flight. In this thesis, data from an eye-tracking device installed in a MH-60S flight simulator was analyzed along with demographic survey data to determine whether correlations between pilot experience, scan techniques and performance level exist under high-speed, low-level flight. The pilots were volunteers from squadrons with Helicopter Combat Wing Pacific, focusing primarily on pilots in a current operational status.

C. HYPOTHESIS

1. Research Questions

- What is the relationship between pilot experience level (in terms of flight hours) and scan techniques?
- When errors are made, such as deviations from assigned parameters, are there any common trends that can be related to a pilot’s experience level? Or are errors as common across all levels of experience?

2. Hypotheses

- 1) H0: There is no association between eye scan pattern (number of fixations, dwell durations, percent time looking out the window (OTW), scan rate between OTW and instrument panel) and flight experience when pilots are classified by total, instrument, or type of mission flight hours.
- HA: With increased flight experience, eye scan patterns will be more efficient: greater number of fixations, shorter dwell durations, faster scan rate between OTW and instrument panel.
- 2) H0: Eye scan parameters (fixation time, dwell time OTW, scan rate) will not predict a pilot’s ability to accurately maintain assigned altitude parameters during a navigation event.
- HA: Eye scan parameters (fixation time, dwell time OTW, scan rate) will predict a pilot’s ability to accurately maintain assigned altitude parameters during a navigation event.
- 3) H0: Eye scan parameters (fixation time, dwell time OTW, scan rate) will not predict occurrence of CFIT.
- HA: Eye scan parameters (fixation time, dwell time OTW, scan rate) will predict occurrence of CFIT.

II. METHODOLOGY

A. PARTICIPANTS

The 17 subjects for the trials were all U.S. Navy helicopter pilots from three squadrons located at Naval Air Station (NAS) North Island, CA. Two different helicopter communities were represented: a carrier-based community, HS (Helicopter Anti-Submarine), and an expeditionary community, HSC (Helicopter Sea Combat). The eight pilots from the HS squadron described themselves as primarily maritime operators. The nine pilots from the HSC squadron described themselves as overland operators. All the pilots, except one, were current in the MH-60S.

Of the 17 participants, 14 were men. The most experienced pilot (3400 hours total) was a female maritime pilot. The least experienced pilot (350 hours total) was a man who was recently certified to fly the MH-60S. Figure 1 shows the distribution of experience level measured in flight hours. From Figure 1, it is shown that the majority of the pilots that participated in the trials were in the 500 to 1000 total flight hour range.

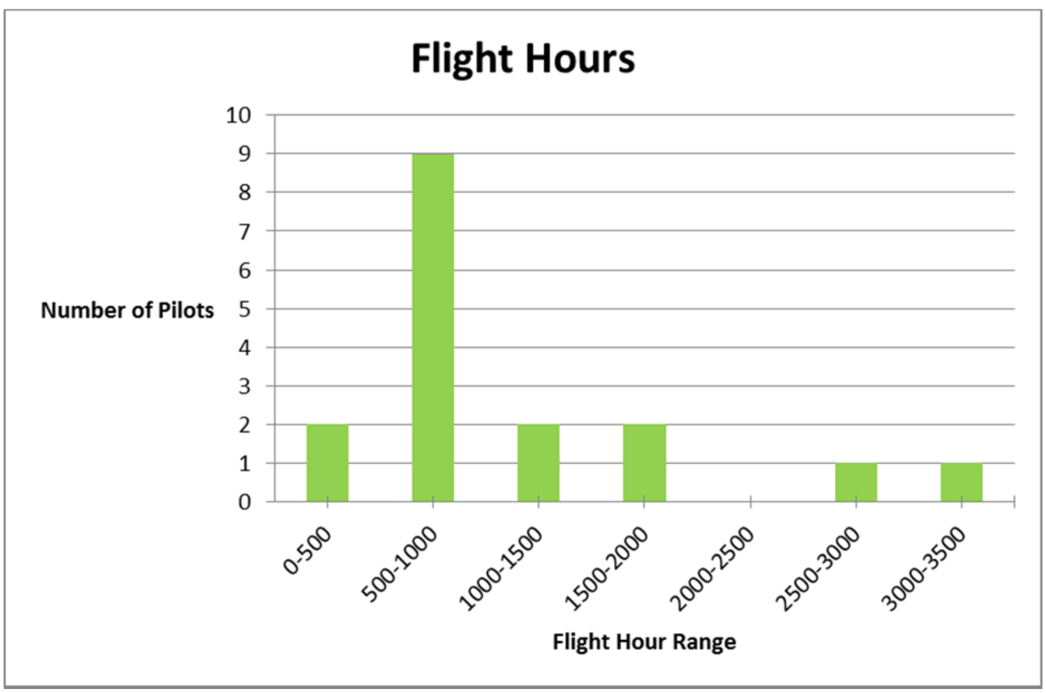


Figure 1. Experience Level, in Flight Hours, of the Participants

Table 1 shows pilot age, experience in terms of years, and how recently they flew a flight and an overland flight. The majority of the pilots were in their first tour, which places them in their middle to late twenties with five to six years of experience. Since all of the pilots were in an operational status, the majority of them had flown fairly recently in relation to the time of the study.

	Age	Experience (Years)	Months Since Last Flight	Months Since Last Overland Flight
Mean	29.70	6.55	0.29	1.64
S.dev	4.09	4.09	0.57	2.76
Min	26	3	0	0
Media	28	5	0	1
Max	40	18	2	12

Table 1. Pilot Age, Experience in Years, and Proficiency

On the flight experience survey, I asked the pilots to estimate the total number of hours they had flown overland. There is no measure of the overland hours in a pilot's log book, and the Navy is just starting to track overland flight time for helicopter pilots as a measure of training. Given that the overland hour data was strongly correlated to the "maritime" or "overland" mission type label (as shown in table 2), I was reasonably certain that most pilots have a good idea of the number of hours they have spent flying overland in their career. Unfortunately, there is no way to verify this portion of the data collection.

	Mean	Std.dev	Min	Median	Max
Land	769.44	367.76	300.00	800.00	1500.00
Mar	409.38	248.73	100.00	362.50	1000.00

Table 2. Comparison of Overland Flight Hours by Mission Type, Land or Maritime

A paired t-test was used to compare the means of the reported overland hours by mission type. The pilots that classified themselves as “overland” had a significantly higher average overland time than pilots classifying themselves as maritime (p = .03046) at the 95% significance level.

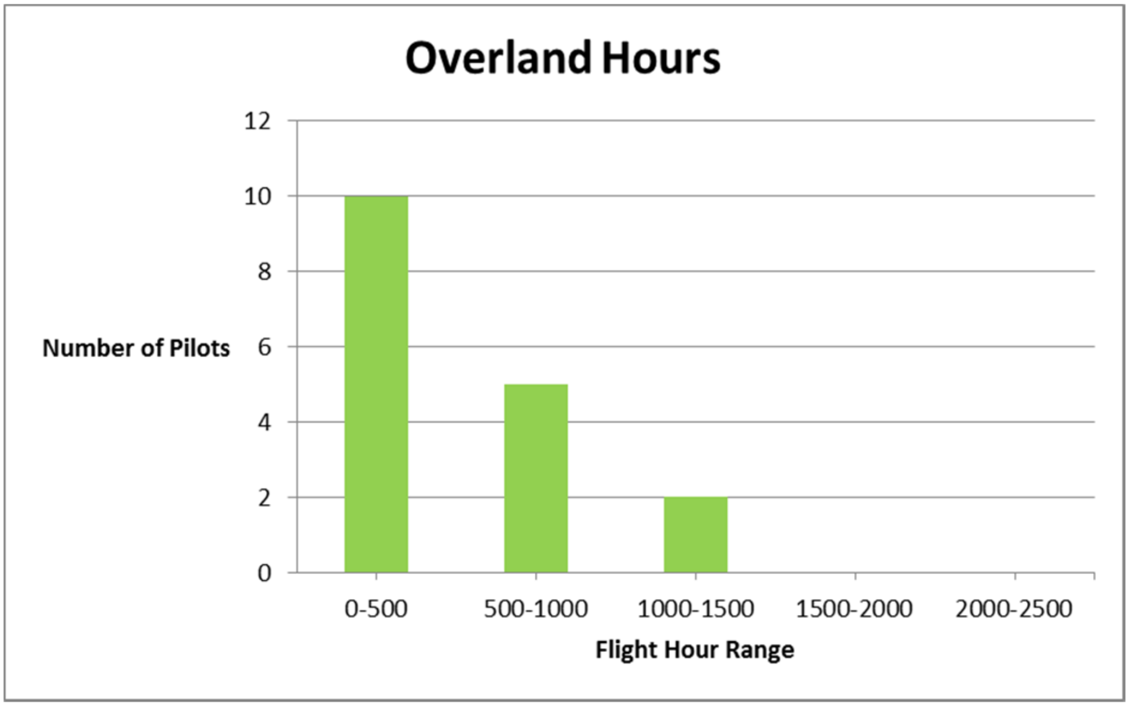


Figure 2. Overland Light Hours of the Subjects, Based on the Report in the Flight Experience Survey

Figure 3 depicts the total flight hours of each pilot against his or her reported overland flight hours. The most experienced pilot in terms of total flight time reported the least amount of overland flight time. This fact was verified through further interview during the debrief process. The subject had spent most of her career in the maritime operational arena and had very little overland flying experience outside flight school. Figure 3’s linear regression formula was

$$y = 0.1689x + 398.82$$

From this data, we can tell a pilot accrues more overland time with experience. This correlation is further shown in Figure 4, in which the pilot who had 3400 hours total

time, but only 100 hours overland time, was removed. This pilot's experience was solely in the maritime environment. In Figure 4 (without the outlier), the linear regression formula was

$$y = 0.4821x + 123.63$$

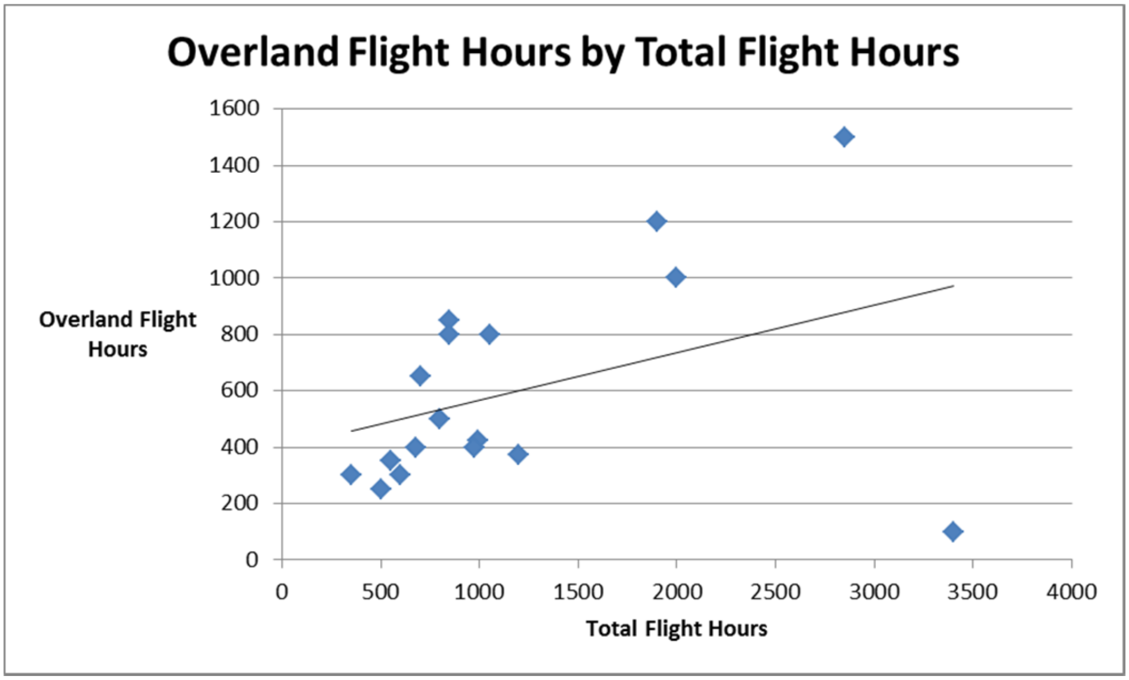


Figure 3. Overland Flight Hours by Total Flight Hours

Figure 4 shows that navy helicopter pilots are spending roughly half of their time flying in an overland environment, with the exception of Subject 8, who only reported 100 hours of overland flying time in her 3400 hour flying career.

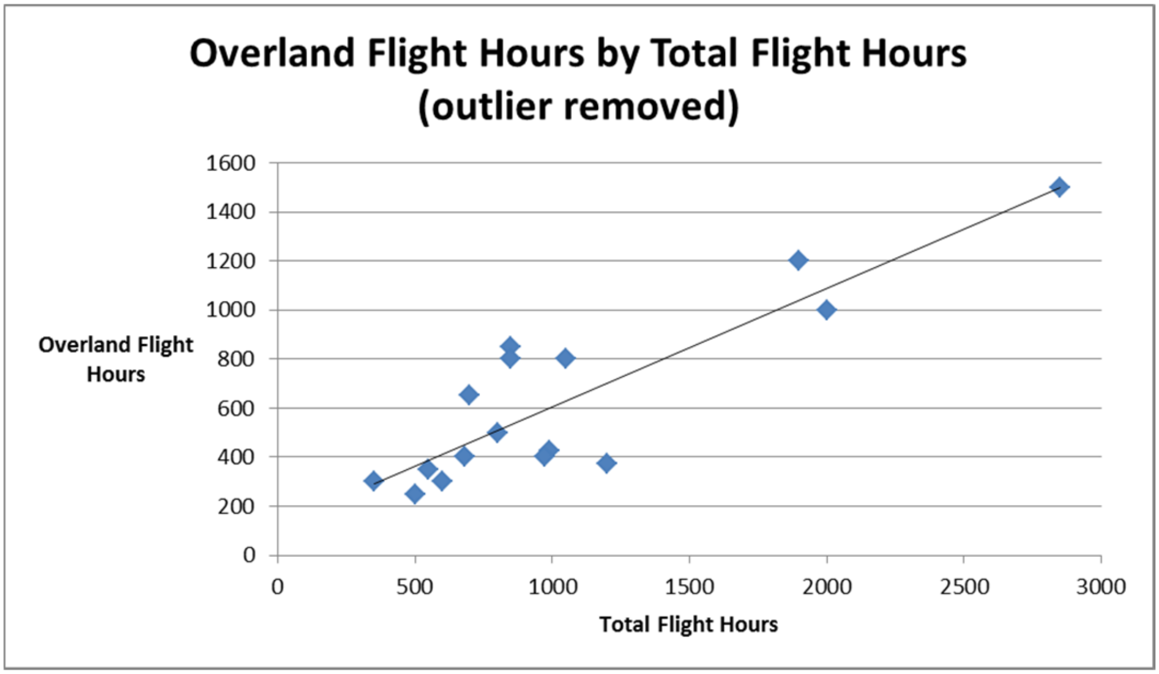


Figure 4. Overland Flight Hours by Total Flight Hours (outlier removed)

B. EQUIPMENT

The experiment was run at NAS North Island using a simulator under the stewardship of the Commander, Helicopter Sea Combat Wing Pacific (CHSCWINGPAC). Permission was obtained for use of the simulator from the CHSCWINGPAC. Helicopter Sea Combat Squadron Three (HSC-3) was responsible for the scheduling of events in the simulator. Scheduling was done through HSC-3.

The simulator was the fixed-base Tactical Operational Flight Trainer 2 (TOFT-2). TOFT-2 accurately represents a MH-60S cockpit (Figure 5). The seat on the right is the flying pilot's seat. The principal investigator, who also acted as a co-pilot, occupied the left seat. Not shown is the simulator operator's chair. This seat was occupied by the research assistant (RA). The FaceLab computers were placed next to this chair, so that the RA could calibrate and run the eye-tracking software. The flight controls and the displays were accurate and current. The TOFT had a full cockpit video system which presented simulated views through the "chin bubbles" which allowed a simulated view of the terrain below the helicopter. Major terrain features, such as roads, mountains, valleys and large buildings were represented accurately in the simulator. Even the Coronado Bay

Bridge, a common sight to every helicopter pilot based at NAS North Island as they fly south for training, was accurately represented in the video display.

Figure 5. The Cockpit of TOFT-2

The simulator collected two types of data through the use of “de-brief” system. This system continuously recorded the simulated aircraft’s status throughout the flight. The first type of data was what the pilot could see—consisting of the aircraft’s location, orientation, airspeed and altitude—were all recorded as each pilot flew the route. This system was commonly used for instruction, but we were able to use it to record flight parameters for the purpose of the study. The second type of data consisted of video recordings. The de-brief system had two cameras: one that recorded the actions of the pilot, and another that recorded the pilot’s flight information data screen. These cameras provide real-time information on what the pilot was looking at during any phase of the flight.

The third source of data was collected by FaceLab. FaceLab, made by Seeing Machines Inc., collects face/head/eye data utilizing infrared light. For this experiment, FaceLab used two pairs of fixed (as opposed to head-mounted) stereo cameras, two infrared light emitters, and two laptop computers. Infrared emitters were needed to produce the level of infrared light necessary for the fixed stereo cameras to capture head and eye motion. Two lap computers ran the software that collected, interpreted, and stored the data from the stereo camera system. Prior to each flight, the FaceLab system had to be calibrated to accurately capture the pilot's head and eye data. To do this a "world" was set up in the computers to simulate the environment in which each pilot operated. A head model was established for each subject in the simulated environment created using the FaceLab software in order to correctly collect gaze and scan data. Once the environment and the head model were established, FaceLab was ready to run through the trials.

Mounting the cameras in the simulator was challenging. The glare shield on the MH-60S cockpit extends high into the cockpit. A position was needed that could both collect the data properly without hampering the pilot's scan as he or she flew the course. The RA used duct tape to mount the top set of cameras, the Out-of-The-Window (OTW) set, far enough back on the glare shield in satisfactory way. The lower set of cameras, the cockpit set, was mounted below the instrument display screens. Only the mechanical "ball" was obstructed. The "ball" is the slip indicator, a ball suspended in liquid that indicates whether or not the aircraft is in balanced flight. During the pilot trials, the subjects informed me that this distraction was not a problem, since the digital "ball" on the display screens was adequate to maintain balanced flight.

The laptop computers were installed next to the FaceLab specialist's operator's chair. This setup allowed the RA to operate both the simulator controls as well as the FaceLab computers. The operator chair was also a great vantage point for the RA to observe each subject as he or she flew. This observation helped greatly when diagnosing problems with the camera's ability to record eye and head movement.

C. FLIGHT SCENARIO

The low-level route the pilots flew in the simulator was created using two VFR Terminal Area Charts of the San Diego area and my personal experience flying in the San Diego area. These charts were used by every pilot from when he or she was in flight school, which made them familiar to pilots of all experience levels. The principal investigator constructed the route by flying it in the simulator, noting the time that each leg required, and checking that the altitude restriction the pilots would have to adhere to was a realistic goal.

The route consisted of ten checkpoints and nine legs, with a total time of about 26 minutes to complete at 100 knots indicated airspeed (IAS). The chart was then marked with course lines and “doghouses”—a popular term for doghouse shaped boxes that align with the legs of the route. Each doghouse consisted of a base heading for the pilot to follow, the length of the leg in nautical miles, and the time to fly the leg at 100 knots IAS. Straight lines were drawn from a checkpoint to its subsequent checkpoint, again offering the pilot a baseline track to follow throughout the flight. The chart was cut down to a manageable size and laminated so that it would survive the many trials to come.

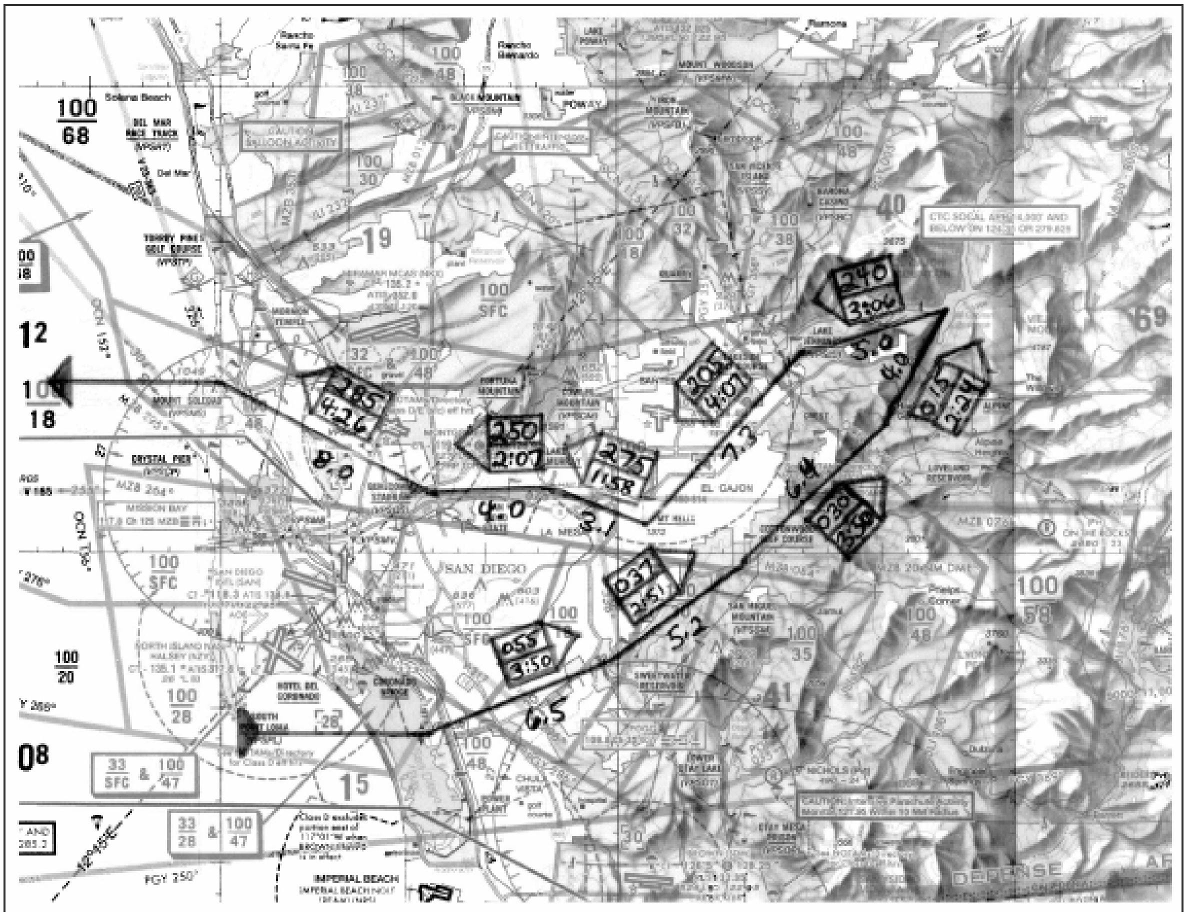


Figure 6. The Chart Used by the Pilot for the Flight

The pilot's task was to safely navigate the helicopter along a route through San Diego's airspace consisting of a number of checkpoints that the pilot identified through either visual means or through the use of instrumentation (in fact, each pilot was expected to back up his or her position using the instruments in the MH-60S). There were multiple places along the route where a CFIT event could occur because the pilots were flying at very low levels among a variety of terrain features. All total, each pilot was in the simulator for no more than one and a half hours each. No adverse conditions were introduced to the flight; that is, the simulated conditions were set to "clear" skies and "calm" winds.

Each pilot was expected to adhere to the following performance parameters:

- Within 100–300 feet above the ground
- Within 10 knots of the assigned airspeed
- Within one mile of the assigned course

Any significant events that occur such as the aircraft's impact with the ground or another object were recorded.

D. SURVEYS

Participants were asked to fill out two surveys as part of the experiment: a pre-flight survey to measure a subject's confidence in his or her ability to navigate, and a post-flight survey to measure the level of difficulty of the flying and navigation tasks. Both surveys were subjective in nature, using a Likert scale of 1–5.

The pre-flight survey consisted of the following questions:

1. To what extent have you participated in activities other than overland navigation that may contribute to improved navigation skills? (Examples may include sport orienteering, land navigation exercises, boy/girl scouts etc.)?
2. At your peak of currency, how would you rate your navigation skills in a low-level (below 200' AGL) overland environment?
3. If tasked today, how would you rate your navigation skills in a low-level (below 200' AGL) overland environment?
4. How much experience do you have with low-level navigation in mountainous desert terrain?
5. How much low-level navigation experience do you have in the Southern California operating area?

Pilots were asked to rate their answers using the following scale:

1. Poor/None
2. Fair/Very Little
3. Average/Somewhat
4. Considerable/Good
5. Extensive/Excellent

The results are shown in Figure 7. Most of the pilots answered the questions in the 3–4 answer range, indicating they were somewhat-to-considerably confident of their ability to navigate the course. The subject pilots also felt somewhat-to-considerably comfortable navigating the mountainous local terrain.

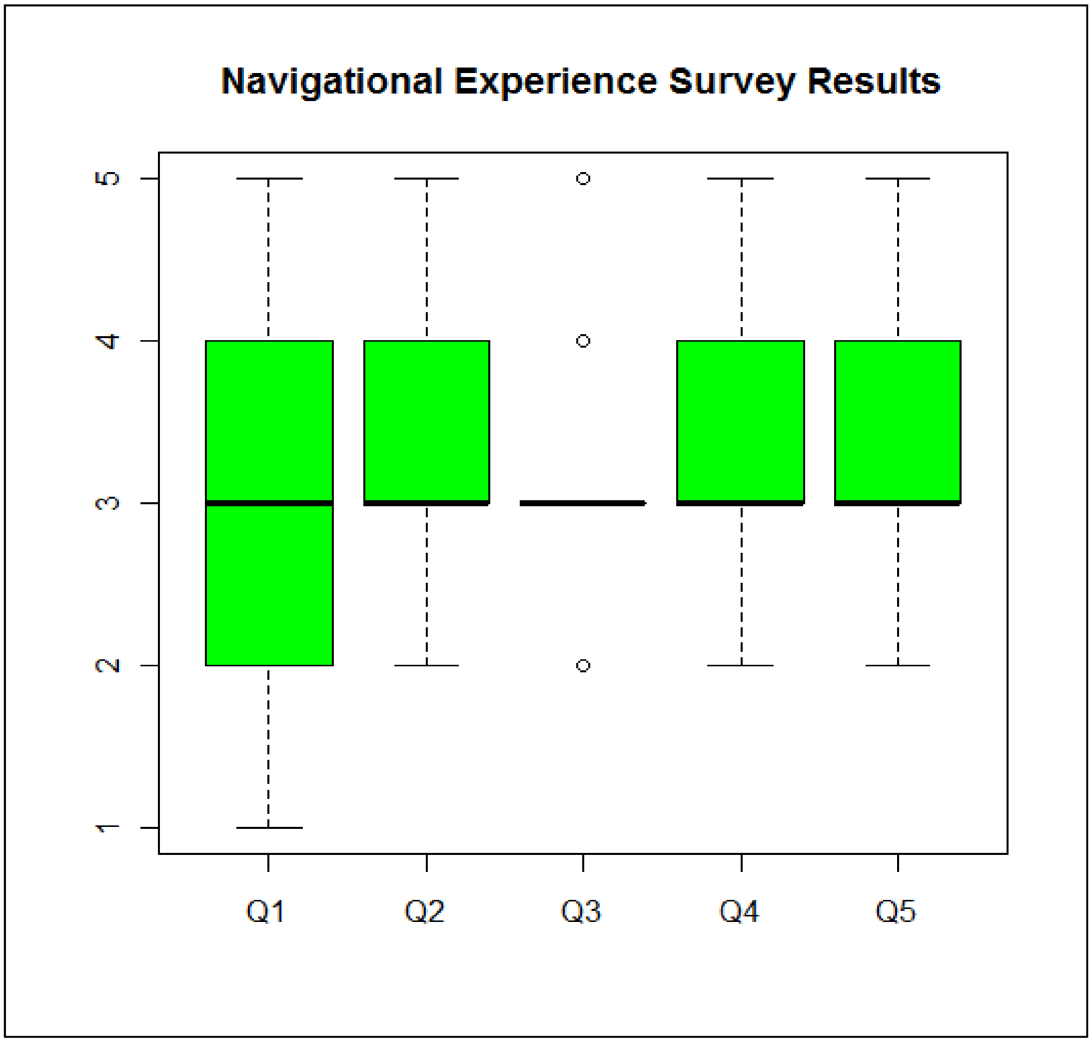


Figure 7. Pre-Flight Navigational Survey Results

After the flight and the FaceLab verification procedures were complete, pilots were asked to proceed to the briefing area in order to fill out a post-flight survey. The purpose of this survey was to gain insight on the perceived level of difficulty for each evolution. The survey consisted of the following questions:

1. How difficult was it to navigate the route while maintaining the assigned parameters?
2. Describe any strategies that you used to stay on course and within the assigned flight parameters.
3. For each navigation leg on the route, please rate how difficult it was to navigate by referencing terrain. Place an “X” on the line that best

describes your experience. No response is necessary for the shaded regions.

4. How confident are you that you flew within the assigned parameters?
5. How confident are you that you correctly navigated the course?

Pilots were asked to categorize their answers according to the following Likert scale:

For questions 1–3:

1. Not at All Difficult/Completely Trivial
2. Somewhat Difficult
3. Moderately Difficult
4. Very Difficult
5. Extremely Difficult/Not at All Possible

For questions 4–5:

1. Very Confident
2. Confident
3. Moderately Confident
4. Not Very Confident
5. Not at All Confident

Figure 8 displays the results of these surveys. All of the pilots except one found the course very easy to fly. Question two was non-quantifiable since the subjects were allowed to write and even expand their answers. Most of the pilots relied on a combination of terrain recognition with the assistance of the co-pilot to navigate the course. Each pilot was asked to rate the difficulty of each leg of the course in question three. Legs three and five were the only legs that seemed have any degree of difficulty associated with them. Most of the pilots were confident they flew the course within the assigned parameters, and they were very confident they navigated correctly in the course.

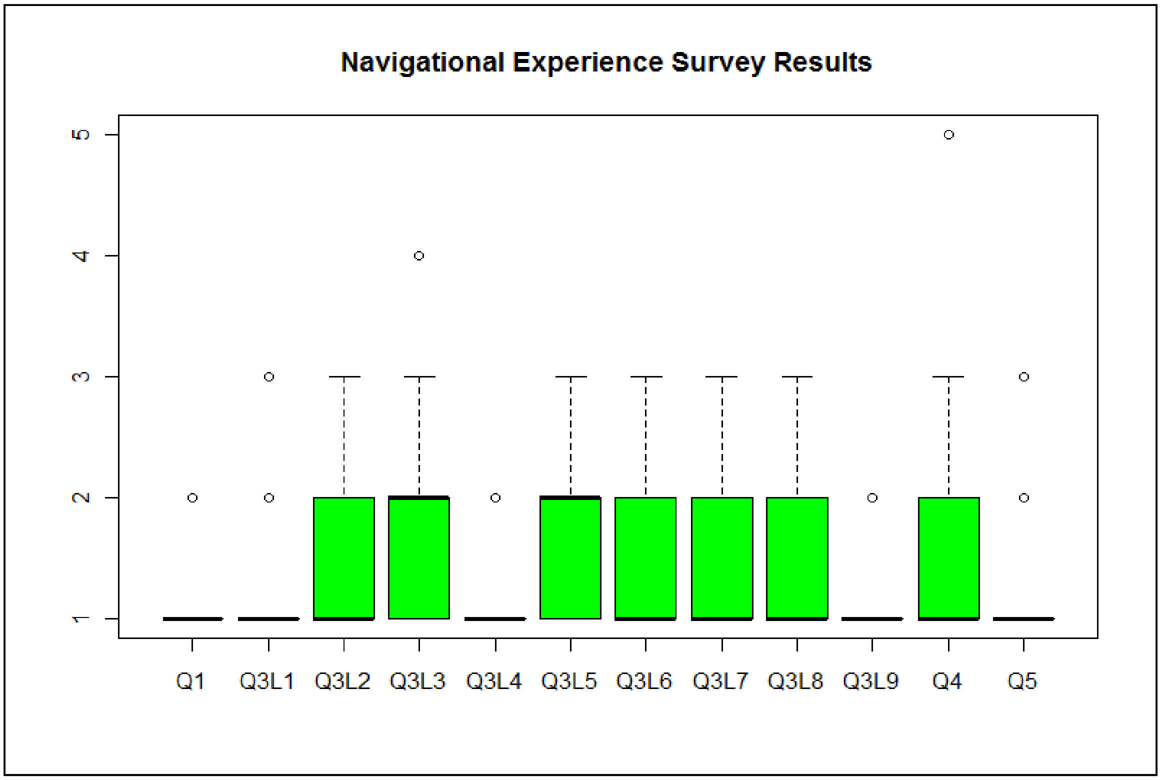


Figure 8. Post-Flight Survey Results

E. PROCEDURE

The procedure consisted of three distinct phases. First, obtaining permission and cooperation from units and their command authority at NAS North Island to use a simulator for testing, and poll the squadrons for volunteers to participate in the trials. Once it was apparent that the use of a simulator and pilot at North Island was achievable, approval had to be obtained from the Naval Postgraduate School in the form of an Internal Review Board (IRB) for testing using human subjects. Finally, after obtaining all approval and cooperation, the team turned to setting up the simulator and conducting the trials in it. This last step had several steps in itself, to include recruitment efforts, a preliminary study, and the trials themselves.

F. STUDY APPROVAL PROCESS

The study approval process began about six months from the desired testing time. Permission was required from the Commander of the Naval Air Forces (COMNAVAIRFOR) Helicopter readiness and from the Commander of the Helicopter

Sea Combat Wing Pacific (COMHSCWP). The permission to conduct experiments in simulators owned by the wing was obtained in writing from the Commodore. Then, coordination between the wing, the squadrons, and the simulation and training device manager had to be made. The squadrons required all experimental runs in the simulator, or “flights” were to be made on a not-to-interfere basis with any training and operations that the participating squadron might be conducting at the time. Timely and constant communication with the training squadron’s operations officer was a necessity in order to ensure that the experiment time blocks remained unmolested. Even though this scheduling was done, there were still several attempts made by some non-participating pilots to “jump” in to time slots already slated for the experiments.

Attempts at setting all of the required coordination through e-mail and phone calls were a failure. In August, four months before the trials were scheduled to begin, the team decided to send the principal investigator (PI) down to make the necessary arrangements in person. Meetings with COMNAVAIRFOR helicopter readiness officer and COMHSCWP chief staff officer both went well, and both seemed enthusiastic about the project. COMHSCWP gave us permission to engage with the wing’s simulation and training device manager (STDM), as well as individual squadrons, to begin setting up the experiment phase of the project. In meeting with the STDM, it was discovered that a MH-60S fixed-based simulator, TOFT-2, was slated for an overhaul in January 2012, and would not be used for regular training during the month of December 2011. Also, the STDM was asked if conducting the experiments in a simulator that is usually used to augment training and operations of over ten squadrons was even feasible, and his answer was yes. The STDM thought it would be possible to use TOFT-2 for three weeks without interruptions and without the risk of someone tampering with the installed FaceLab equipment. As long as HSC-3 allowed it, our team had full use of a simulator for three weeks, with no cost.

Three more meetings were held on this first trip: with the commanding officer of HSC-21, with the executive officer of HSC-3, and with the executive officer of HS-4. All were very positive and enthusiastic about supporting the project. Subsequent meetings were held with the operations officers of all three squadrons to do the following:

- An introduction, and an exchange of contact information
- A description of the project concept and goals
- A description of the project's needs in terms of pilots and time commitment

These points were outlined to the operations officers, with no initial commitments made. Due to the dynamic nature of helicopter operations, none of the operations officers could say with certainty, four months prior to the scheduled experimentation period, that they would be able to commit time and pilots to the project.

After the first trip, e-mail contact with the above mentioned was established and the Internal Review Board (IRB) process was initiated. Two more trips down to NAS North Island were scheduled in order to set up the simulator and confirm the availability of test subjects from the squadrons. The team began to make travel plans for the investigator and the RA for the three weeks we would need to conduct the experiments using TOFT-2. Also during this time, the team clearly defined the goals for the simulator use with the FaceLab equipment.

G. EXPERIMENTAL SET-UP PROCEDURE

During the last week of November, the PI left NPS for NAS North Island to begin setting up the actual workspace and recruiting subjects for the trials. The FaceLab equipment, consisting of three metal briefcases, was loaded in to a rental car and transported to San Diego. Upon arrival, contact was made with COMHSCWP's Simulation and Training Device Manager (STDM) to get the status of TOFT-2. He and his team were trying to fix the video system on the co-pilot's display and were still trying to get the de-brief system up and running. The STDM and his team were sure the video would be running in time for our trials, but he was not so sure about the de-brief system. Next, the simulator space was inspected to get an idea of the working environment. The research assistant and the investigator had a briefing area to work in, which served as a good place to meet with subjects before and after the trials to fill out consent and survey forms, as well as give them a description of what the trial consisted of. There were plenty of power outlets both in the cockpit area of the simulator and outside for the FaceLab laptops to plug in to. Lastly, it was noted that there were two computer terminals we

could use that had internet access just in case. The highly qualified research assistant informed me that these would be necessary to upload the data to NPS every night.

H. RECRUITMENT PROCEDURES

The next task was to go to each of the participating squadrons, give them a brief on the project, and obtain contact information to start populating a subject list. The personnel were briefed on the goals of the project, what their participation consisted of, and any benefits, risks or hazards associated with the trials. Sign-up sheets were passed around and returned to the investigator via a sealed envelope if they were filled out at that time. Each participant also had the option to fill out the sheet at a later time and place them inside a folder to be picked up a day later. On these sheets, participants gave their contact information and their preference of flight time and day. Again, the project was well received by the squadrons and the initial pledge of support was overwhelming.

That week, solicitations were sent through e-mail for volunteers. Each person who filled out a contact sheet was sent an introductory e-mail with a time slot for their flight. Each e-mail was sent to the individual only, and no reply was required. A schedule was created on a spreadsheet consisting of a subject identification number and the time and date of their flight. Each day consisted of three available time slots of one and a half hours each. The team was guaranteed use of TOFT-2 from 1200–1700 every day, so the three slots fell within that window.

The last Friday of every week was used as a recruitment day. A squadron was selected, and briefed on what their participation would entail, as well as the possible benefits to rotary-winged aviation as a whole. Throughout the brief, the voluntary nature of the experiment was stressed, and that the research team had no affiliation with their command in any way. Contact was also made with the COMHSCWINGPAC offices and Mr. Jacobs to give updates on the status of the simulator and the project. The support given by the wing staff and the simulator managers was top-notch. At the end of every Friday, recruitment e-mails were sent to the volunteers from the squadrons to confirm the time of their flights.

I. PRELIMINARY DATA COLLECTION

Week two began with approval from the IRB to start conducting our research. COMHSCWINGPAC operations was approached in order to solicit volunteers for the preliminary (pilot) study phase of our research. These pilots tended to be more experienced, and they would offer good insight as to how to integrate FaceLab into a flight with as little interference as possible. The STDM informed the team on Monday morning of our second week that the de-brief system, to include the video recording system, was working. The RA arrived from Monterey that morning to assist with the set-up of FaceLab and to run the software during the pilot trials. The RA set up the world, the computer representation of the cockpit, in the FaceLab software. In order to do the set up, he had to use cardboard cut-outs and duct tape to create “maps” of the out of the window (OTW), instrument and helicopter diagnostic displays. The investigator served as the first head and eye model for creation in this new environment. Through much trial and error, the RA was able to establish the vertical link between the upper and lower eye-tracking cameras on the first day. The first two complete test runs of the route were flown in conjunction with the FaceLab system. A few things were brought to light on these first test runs. Certain actions by the pilot, such as resetting the aircraft’s timer, interfered with the eye-tracking cameras. During these initial runs, we found that the total time to complete the route at 100 knots indicated airspeed (IAS) was 26 minutes and 30 seconds.

On the first day of preliminary trials, we were able to get two COMHSCWP pilots though the first pilot studies while successfully collecting data. There were a few problems that were overcome. The pilots both stated that they would prefer to have a co-pilot (the investigator) navigate for them; that is, they wanted someone in the left seat reading the chart, relaying important information to them. This was contrary to our first ideas about the design of the study, in which the pilots would fly without any help from a co-pilot, referring to the chart which would be taped over the multi-function display (MFD) used to relay the status of the aircraft. Both pilots on the first day of the pilot trials stated that this setup was not realistic, in that they would not fly that way, and they found it distracting since it was not what they were used to. Because both pilots relied on the co-pilot for navigation, the previously noted problem with the timer use was no longer a

problem since the flying pilot was not the one resetting the timer. It was also noted that both pilots tried to adhere to the published "course rules" of the area, in that they tried to maintain the parameters outlined in the local area flying rules dictated by NASNI flight operations. In both cases, they were told that this was not necessary. For the purposes of the research, it would be better if they tried to maintain the altitude restrictions delineated in the brief. Both pilots were very concerned with the presence of power lines on the route. They were told that none were present in the simulation, and it was noted that this fact needed to be included in the pre-flight brief. The pilots also stated that they flew a constant torque when flying low-level over terrain not a constant airspeed. This phrase means that the helicopter is flown at a constant power setting and is allowed to increase or decrease speed, depending on the attitude needed to maneuver over the terrain. Despite this difference, the 100 KIAS approximation used earlier was still valid, in that the average speed over the flight was 111 KIAS (standard deviation 8.5).

On the second day of preliminary trials, two more pilots from COMHSCWP flew the simulator with FaceLab installed. During these trials, the team tackled the issue of "pilot creep"—that is the tendency for pilots to crouch forward and lower as the flight progresses. The RA suggested that these two pilots use the shoulder harnesses while flying the simulation as a way to limit the pilot creep, and thereby eliminating the need to re-adjust the eye-tracking cameras after the first twenty minutes of every flight. This procedure was not welcomed by the pilots. The simulator is a non-motion simulator and the harnesses are not normally used when flying TOFT-2. Aside from some grumbling, all pilots in the trials complied with this request, and the technique proved successful in reducing the number of flight stoppages due to the loss of eye-tracking "lock."

J. PROCEDURES DURING TESTING

With the preliminary trials complete and the procedures modified to reflect the findings of those test runs, it was time to move on to the testing phase of the research. TOFT-2 was reserved for our use Monday–Thursday of every week, for the hours of 1200–1800. Three trial runs, or flights, would fit within the allotted time slot. This allowed adequate time for the pilot briefing, set-up of the FaceLab model for each pilot, the flight, and the de-brief.

Upon arrival, each pilot was greeted by the primary investigator (PI) and shown to the designated briefing area. The briefing area used was the same briefing area used for flight training, so the place was familiar to every MH-60S pilot. A brief outline of the flight and what would be required of the pilot was given by the PI, to include an overview of the route and the parameters of flight expected they were expected to maintain. During the route brief, the pilots were told not to adhere to the normal course rules for the San Diego area, and to maintain the altitudes described in the Flight Scenario section. The pilots were also told that the route in the simulator contained no hazards to low-level flight, such as power lines or other low-flying aircraft. During the initial briefing, pilots filled out the consent to participate in research from, the flight demographics survey and the navigation experience survey. Once the briefing and the surveys were completed, the pilot and the PI would then join the research assistant in the simulator. The research assistant stayed in the simulator between flights to finish saving the data from the last flight, and to set up the FaceLab system for the next flight.

The RA took over at this point. He was responsible for making sure each pilot was strapped in to the seat, to include the use of the shoulder straps. Once the pilot was strapped in and had his or her seat properly adjusted, the RA would begin to build the face and eye model in FaceLab. Each pilot was asked to track his or her eyes across the cardboard cutouts created by the RA in order to calibrate FaceLab to their gaze. The majority of pilots would have to go through the entire set-up process, which could take up to 45 minutes. In a few cases, the pilot currently occupying the seat was close enough to height and seat adjustment to the prior pilot. These cases allowed the RA to do an abbreviated set-up procedure, which only took 15 minutes. Either way, once the pilot's face and eye gaze were set up in FaceLab, it was time to start the flight.

Once the set-up was complete, the pilot was instructed to take off from pad three at NAS North Island and start flying towards the starting point of the low-level route (checkpoint one), commonly known to all San Diego-based pilots as the "Blue Crane." During this time, the PI stepped out of the simulator to start the recording of the flight using the simulator's de-brief system. By the time the PI returned from this task, each pilot was usually at the Coronado Bay Bridge. Each pilot flew the course rules altitude

until reaching the area of checkpoint one, at which they descended down to the 100–300 feet AGL requirement of the trial. At each checkpoint, the PI relayed the recommended heading to the next checkpoint, the time of flight between checkpoints at 100 KIAS, and any significant terrain features the pilots could use to find the next checkpoint. The PI also informed the RA which checkpoint we were at, so he could mark the point in the FaceLab database. Throughout the flight, the PI notified the pilot if they were in violation of any of the assigned parameters, but this notification not given often.

Upon reaching the tenth and final checkpoint of the flight, the PI “froze” the simulator, suspending any further flight. The PI crawled out of the cockpit to stop the de-brief system’s recording. The RA went through FaceLab’s validation process, in which the pilot’s gaze and position relative to the simulated “world” created by the software was verified. After the validation was complete, the pilot was asked to step out and return to the briefing space. The total flying time in the simulator for each pilot was a maximum of one half hour. The set up time and validation time totaled one hour. The pilots filled out the post-flight survey, complete with comments on how to improve the simulation. After the survey was complete, the pilots were given some time to ask questions about the experiment and what we were trying to do, since they had completed the scenario and there was no longer any risk of this information tainting the results.

K. PROBLEMS ENCOUNTERED DURING TESTING

There were a few problems encountered during the testing phase of the thesis. Previously mentioned in the preliminary data collection, the “pilot creep” kept causing problems with camera alignment. The RA’s suggestion of having the pilots use the shoulder straps helped keep the pilots from slumping, but the added procedure did not eliminate the problem entirely. Many pilots would not only “sink” into their seats as they flew, but would also lean to the left as they used the collective flight control input to adjust the helicopter’s power. The only solution to this problem was to freeze the simulator and re-adjust the cameras to account for the pilot’s adjusted position. “Freezing” a flight is unrealistic and distracting, and elimination of this problem would only increase the reality of each flight.

Another problem encountered was the tendency for some pilots to be “jittery” during the flight. A few of the pilots showed what both the RA and the PI called “excessive” head movements during the flight. During the post-flight debrief, the pilots with this problem were not asked if they were hyper-aware of their head and eye motion due to the presence of the cameras. The RA and the PI both thought this might be the case, but no way to verify this was achieved during the scope of this experimentation. With these pilots the cameras were unable to track their head and eye movements. The result was degraded but still usable data from these pilots.

The time to set up the cameras for FaceLab seemed to be directly related to the height difference from one pilot to the next. Larger height differences from pilot-to-pilot resulted in significantly longer FaceLab set-up times. The RA and the PI could not think of any easy way to get around this difficulty. One possible resolution is to have set camera positions based on the height of the pilot. This was not possible to do in the short amount of time given to execute the experiment, but setting adjustments based on height may save time in future efforts using FaceLab.

There was a tendency for at least one pilot a day not to show up for the scheduled flight. This was the price of conducting trials in the field. The flight operations each squadron conducts are at the whim of weather and aircraft availability. There are myriad reasons a pilot could have missed the trial obligation. An aircraft’s suddenly becoming available for a flight event, or flight delays due to weather are just two examples of why a pilot might be unable to meet the scheduled trial commitment.

One of the squadrons insisted on the flight events being placed on their daily flight schedule. This did convey the commander’s intent to support the project, but the flight schedule is a signed order from the commanding officer of the squadron to all that are placed on it. While the intent of the commanding officer was to be helpful, they were informed that this violated the voluntary nature of participation in the study. To avoid any further command involvement, the scheduling e-mails were sent directly to the participants by the PI. All further events from that point on were only blocked out on the squadron’s flight schedule as “NPS Experiment–TOFT-2” with no further information or identification of the participants.

L. EXPERIMENT PHASE CONCLUSION

Upon the conclusion of the experiments, the FaceLab system was dismantled and removed it from TOFT-2. All the items were inventoried and packed for the trip back to NPS. The simulator was cleaned out and returned to its original operating condition. All TOFT-2 de-brief files relating to our experiments were deleted from the simulator's database. All written materials and files, including the hard drive, were locked in briefcases by both the RA and the PI. These were not opened again until they reached the lab at NPS.

The next day, the PI visited the participating squadrons, the wing and the training and simulation device manager, giving them each a wrap up of the experimentation phase and a few words on the next steps in the project. They were all thanked for their overwhelming support. The PI returned to NPS with the FaceLab equipment and the data.

III. ANALYSIS AND RESULTS

A. DATA PREPARATION

1. Flight Briefing and Surveys

The data came from three sources: the surveys that the subjects filled out before and after the flight, the output from the simulator's debrief system, and the output from the FaceLab eye tracking software. Each of the data sets had to be handled in a unique way, due to its format from the source. The surveys were in a pen and paper format, while the debrief data and the FaceLab data were in the form of comma-separated files (CSV) saved on a secure hard drive.

2. Pre-Flight Briefing and Surveys

During the initial briefing, the pilots filled out the consent to participate in research form, the flight demographics survey and the navigation experience survey. The results from these surveys were translated from paper to an Microsoft Excel spreadsheet for each pilot. Flight demographics and navigation experience surveys were combined into a single spreadsheet to give a picture of pilot experience. The data were saved for comparison later with the de-brief and the FaceLab data.

3. Simulator Data

During each trial run, the simulator's de-brief system recorded the parameters of flight for the simulator. These included the simulated aircraft's airspeed, altitude, heading, and data specific to aircraft systems, such as the status of the hydraulic systems. There were over 50 recorded parameters in each file pertaining to each subject's flight. They were recorded over a series of time frames. At any given time mark, the de-brief system recorded the aircraft's performance and position. For the purpose of the study, only 11 of the parameters were needed. They were:

- The magnetic heading (degrees)
- The aircraft's position in latitude and longitude (2 parameters) (degrees)
- Whether or not an instance of an aircraft "crash" had been reported (0 or 1)

- The helicopter’s ground speed (GS, in knots)
- The helicopter’s indicated air speed (IAS, in knots)
- The aircraft’s pressure altitude (or reading on the pilot’s barometric altimeter, in feet)
- The aircraft’s height above the terrain (or the reading on the pilot’s radar altimeter, in feet)
- The helicopter’s vertical speed, up or down, in feet per minute (feet per minute)
- The position of the pilot’s “ball” or sideslip indicator, the instrument that indicates whether or not the helicopter is in balanced flight (degrees)
- An indication of the actual sideslip of the aircraft (degrees)

Out of the data files for the 17 subjects, only two were corrupted. The technicians responsible for restoring the de-brief system during our pilot trials issued a warning that corruption might happen. The de-brief system had a tendency to get “stuck” on one data point and remain there for the duration of the flight. There was no way to recover this data other than re-flying the flight and conducting another trial. Unfortunately, this option was not realistic for many reasons. Most of the pilots were already taking time out of busy schedules to accommodate these trials. Also, a pilot would fly differently on his or her second run than on the first. Fortunately, the corruption only occurred for two of the flights, so most of the data was available for analysis.

Each file contained data on the entire route of simulated flight, from when the pilot took off from pad three at Naval Air Station North Island until the simulator was frozen at the end of the trial run. For the purposes of the study, only data from the route of flight depicted on the chart was used. Each file was trimmed so that only data along the route of flight from checkpoint one to checkpoint ten was included. This editing was done by noting the heading change when each pilot turned off the ingress route to intercept the course and checkpoint one, and again noting the heading change at the end of the route at checkpoint ten.

4. Eye Tracking/FaceLab Data

The FaceLab data came in three distinct files for each participant:

- Timing Files: Each checkpoint was marked to a corresponding frame number.
- World-View Files: The data indicating where the pilot was looking at a given frame number.
- Eye Files: Data on the pilot's gaze and saccades. These were not used in the analysis for this thesis.

Each file had to be processed in order to obtain useable data for hypothesis testing. This processing was done using the open-sourced statistical software [R], "a powerful tool for statistics, graphics, and statistical programming" (Teeter, 2011). For each file, a loop was written and executed to gather the data needed for further analysis. The [R] code used is presented in the appendices.

5. Timing Files

The timing files were processed first in order to attain the checkpoints and their corresponding frame numbers. The raw data from the FaceLab software was in the form of text files (.txt files), which had to be converted to comma-separated files (.csv files) for use in [R]. The output files were in CSV form and imported in to Microsoft Excel for analysis. This process was common to all of the files produced by FaceLab. Table 9 shows a typical result after the processing of the timing files:

TIMING					
Min.		Median	Mean		Max.
13600		84540	84580		156300
FRAME_NUM	EXPERIMENT_TIME	GMT_S			Check pt
42362	706.117	1.32E+09			1
55579	926.384	1.32E+09			2
66751	1112.57	1.32E+09			3
80485	1341.45	1.32E+09			4
87289	1454.84	1.32E+09			5
98357	1639.29	1.32E+09			6
110451	1840.84	1.32E+09			7
115429	1923.8	1.32E+09			8
123824	2063.71	1.32E+09			9
135408	2256.76	1.32E+09			10

Table 3. A Sample of the Timing Data for Subject 7

The minimum and maximum frame numbers represent the beginning and the end of each flight from take off until the simulator was frozen after checkpoint ten. Flight time duration for each subject was calculated using the experiment time (in seconds) from checkpoint one to checkpoint ten. For subject number seven, the flight time from checkpoint one to checkpoint ten was 1550 sec, 25 minutes and 50 seconds.

6. World View Files

Timing data with the checkpoint locations in terms of frame number and mission time was critical for further analysis on the world-view files. Only data between the start of the flight at checkpoint one and the end of the flight at checkpoint ten was necessary. Using the frame numbers corresponding to the checkpoints in the timing data, each world-view file was cropped so that only the information in between these checkpoints was included in the analysis. This step was critical because the focus of the study is on cruise flight only. Data before checkpoint one includes the take-off from the pad on NAS North Island and the over flight of the Coronado Bridge, which would require altitudes greater than 500 feet above ground level. The cropping process was completed using the similar file conversion (text file to CSV file to Microsoft Excel, and then back to CSV) as used for the timing files.

Once each world-view file held the correct segment of data and was in the correct format, [R] was used once again to process the files to get the necessary information for analysis. Each subject’s world-view file contained a column called ITEM_NAME, depicting the item the pilot was looking at during the time stamp. The items were:

- InstrumentDisplayCockpit
- InstrumentDisplayOTW
- MapCockpit
- MapOTW
- Nothing
- OTW
- OTWCockpit

The “Nothing” category refers to data that was not captured by the recording system.

The first item in the item name indicated the actual item being scanned. The second part of the name indicated that the camera that was used to pick up the item being scanned by the pilot at that particular time. For example:

InstrumentDisplayCockpit

The item being scanned at the frame

The camera used to pick up the scan

The item “OTW” does not have a second identifier. The OTW camera picks up this item, and it is assumed that the designers of FaceLab thought the identifier “OTWOTW” would be a little redundant.

Again, using [R], each subject’s file was processed in order to ascertain how many of each item was present in the ITEM_NAME column. At each frame number, the pilot was looking at one of these objects. The number of times an item was picked up corresponded to the number of frames the pilot was looking at that object. After running the code, [R] produced a CSV file with the desired output.

The [R] code (see appendix B) produced the following output (shown after the CSV file was imported into Microsoft Excel):

Instrument DisplayC	Instrument DisplayO	Map Cockpit	Map OTW	Nothin g	OTW	OTW Cockpit
37296	356	453	1	45300	8800	644
Frame number	Change from	To				
42408	1	5				
42450	5	1				
42473	1	5				
42496	5	1				
42507	1	5				
42508	5	1				
42510	1	5				
42616	5	1				
42627	1	5				
42635	5	1				
42660	1	5				
42763	5	1				

Table 4. A Sample of the World-View Data After Processing in [R] and Microsoft Excel

The top row depicts which object the pilot was scanning at the time. The second row shows the number of "hits" for each object. The third row delineates the frame number at which a pilot shifted scan from one object to another. For example, in the highlighted row, the pilot shifted his or her scan from object 5 to object 1, corresponding to a shift from the instrument display (picked up by the cockpit cameras) to nothing.

Since the only concern was what the pilot was looking at during a certain time, there was no need to discriminate between which camera picked up the information. By adding some additional columns, the object's camera location was eliminated from the analysis. This produced a column that only indicated which object the pilot was looking at during a given frame number. Additionally, a "diff" column was added to indicate the number of frames that passed between the pilot's last scan shift.

Instrument DisplayC	Instrument DisplayO	Map Cockpit	Map OTW	Nothing	OTW	OTW Cockpit
37296	356	453	1	45300	8800	644
Frame number	Change from	To	Diff	New	places	
42408	1	5		1	4	
42450	5	1	42	4	1	
42473	1	5	23	1	4	
42496	5	1	23	4	1	
42507	1	5	11	1	4	
42508	5	1	1	4	1	
42510	1	5	2	1	4	
42616	5	1	106	4	1	
42627	1	5	11	1	4	
42635	5	1	8	4	1	
42660	1	5	25	1	4	
42763	5	1	103	4	1	

Table 5. World-View Data with the Frame Difference Column and the Combined Objects

Figure 9 shows a guide that was created in each spreadsheet in order to translate the [R] code:

Old Code	
1	InstDispC
2	InstDispO
3	MAPcp
4	MAPotw
5	Nothing
6	OTW
7	OTWcpt

New Code	
1	Inst_Disp
2	MAP
3	OTW
4	Nothing

Figure 9. Code Descriptions, Showing Numbers Corresponding to the Scanned Object

Once the data from the world-view and timing files were processed in Microsoft Excel the results were combined on to a final page in the workbook for further analysis. Each data type was then combined on a final worksheet for comparison along pilot

experience level and type. When the data from the scan location, or where exactly the pilot was looking at a certain time was put together, it became apparent that the "Nothing" category was a problem. The "Nothing" category refers to data that was not captured by the recording system.

	Dwell	%		
Subj #	Inst_Disp	MAP	OTW	Nothing
7	40.55143	0.488961	10.17124	48.78837
8	8.00052	57.66396	3.373641	30.96188
11	43.62472	0.224121	42.82276	13.3284
12	37.03034	1.062985	44.38704	17.51964
13	1.330562	0.186013	64.95838	33.52504
14	3.380094	0.259408	26.01536	70.34513
15	33.95664	0.230613	41.51316	24.29959
16	7.187941	41.78071	26.34036	24.69099
18	18.44136	0.25168	39.68335	41.62361
19	26.78608	0.074518	39.00145	34.13794
20	4.515534	1.125352	6.021339	88.33778
21	3.498846	6.829686	23.64647	66.02499
22	5.056483	0.427872	72.02139	22.49425
23	6.997912	0.254952	45.97211	46.77503
25	6.131998	0.120843	26.58025	67.16691
26	36.48841	0.108766	14.71449	48.68833
28	11.86348	0.441894	43.45943	44.2352

Table 6. Percentages of the Total Flight Time a Pilot was Looking at a Certain Object

In order to accurately represent where the pilot was looking at a time in which the FaceLab system recorded "Nothing," a substitution had to be created based on the flight videos from the simulator's de-brief system and the videos based on the world-view recreation created by the FaceLab system. FaceLab, using a world-view program, was able to re-create a simulation that depicts the location of the pilot's eye gaze, as well as the orientation of their face at any given time. The de-brief videos were watched to ensure that, at no time during the flight, the pilots were looking at "Nothing." They were indeed looking at something, whether it was out of the window or down at the instrument

or map display. Next, the world-view videos were watched to see what FaceLab thought the pilot was looking at a given time. The insights gained from these viewings produced an interesting result.

During the set-up of the simulated “world” in FaceLab, which took place in TOFT-2, the RA measured the instrument view screen and entered the dimensions (.28 x .18 meters) in to the FaceLab simulation. He also measured the distances from the top of the glare shield and the space in between the map and instrument displays to get an accurate representation of the cockpit display. While his efforts are commendable, the world-view videos show that the alignment of these displays may have been in error. The videos clearly show that the pilots are scanning down towards the instrument display, yet, because the display in the world-view simulation was so small, the scan often “missed” according to FaceLab, and the event was recorded as if the pilot was looking at “Nothing.” Only in three cases, subjects 14, 20, and 25, did “Nothing” actually mean “Nothing.” In these cases, the FaceLab system was actually unable to register the location of the pilot’s eye gaze or head model location. The simulation was unable to reproduce an accurate representation due to poor gaze quality, the calibration of the system was off, or because the pilot made a movement that the system was not able to pick up. The percentages of “Nothing” in these cases represented a loss of data recording.

Armed with this new realization, the data was modified in the Microsoft Excel spreadsheet to accurately represent scan location. The “Nothing” category was switched to “Inst_Disp,” and the percentages were re-calculated. Each subject’s world-view sheet was modified separately in order to ensure that the findings of the video review were accurately represented. The combined data checking resulted in the percentages and times used in the data analysis in the next chapter.

	Adjusted	Dwell	%
Subj #	Inst_Disp	MAP	OTW
7	89.33	0.48	10.17
8	38.96	57.6	3.37
11	56.95	0.22	42.82
12	54.54	1.06	44.38
13	34.85	0.18	64.95
14	73.72	0.25	26.01
15	58.25	0.23	41.51
16	31.87	41.7	26.34
18	60.06	0.25	39.68
19	60.92	0.07	39.00
20	92.85	1.12	6.021
21	69.52	6.82	23.64
22	27.55	0.42	72.02
23	53.77	0.25	45.97
25	73.29	0.12	26.58
26	85.17	0.10	14.71
28	56.09	0.44	43.45

Table 7. Corrected Percentages of the Total Flight Time a Pilot is Looking at a Certain Object

Using the results from the [R] code combined with some coding in Microsoft Excel, a metric was developed that gave some insight as to the direction of a pilot's scan. The Microsoft Excel code was based on the "change from" and "to" columns. A count was taken to determine how many times a pilot switched his or her gaze from one object to another. For example, if a pilot switched his or her gaze from the instrument display to out-of-the-window, it would be recorded as an event of that type. Using this method, the following table was produced for further analysis:

		Scan	Direction		
IM	IO	MI	MO	OI	OM
110	932	111	1	931	2
346	1355	339	34	1361	28
41	1116	40	2	1118	1
71	1213	75	0	1208	4
37	1259	43	6	1253	12
42	2569	49	12	2563	19
29	1308	30	3	1308	4
75	3529	62	27	3541	14
50	2784	52	3	2782	5
13	1864	13	3	1864	3
237	853	239	8	851	10
318	1526	316	99	1529	96
35	1571	36	4	1571	5
81	2231	74	21	2237	14
62	3243	64	7	3242	9
19	1581	18	3	1582	2
105	2759	103	31	2760	29

	Scan	
Code	From	To
IM	Inst_Disp	MAP
IO	Inst_Disp	OTW
MI	MAP	Inst_Disp
MO	MAP	OTW
OI	OTW	Inst_Disp
OM	OTW	MAP

Table 8. Occurrences of Scan Shifts with the Codes Used for each Type of Event. One Row for Each Pilot

Each column represents a scan shift event. When a pilot transitioned from scanning the instrument display to out-of-the-window, that event was recorded as an "IO." The first row represents the data collected for Subject 7. In this example, Subject 7 shifted from the instrument display to out-of-the-window 932 times during the flight.

This concluded the data set up for analysis. The eye-tracking data collected from the experimental trials in TOFT-2 will be used in later analysis for future research questions.

B. STATISTICAL ANALYSIS USED

Trend analyses were done using the measures of rank correlation outlined in Conover's *Practical Nonparametric Statistics* (1999). The type of test for trends used was Spearman's *Rho* (Conover, 314). The tests used the data based on bivariate samples to see if a trend existed as the one sample is ranked according to the other. A negative

Rho value indicated an inverse relationship; a positive value indicates a direct

relationship. For example, a negative *Rho* value indicated that altitude standard deviation decreased as pilot experience increased.

Helicopter group analyses utilized a two sample t-test assuming unequal variances and the scan direction data was analyzed using paired t-tests in [R].

C. DATA ANALYSIS AND RESULTS

1. Pilot Performance Analysis

The debrief data from TOFT-2 provided the performance data for each subject's flight. Using this data, Table 9 is presented:

	Mean	Median	Stdev
Time to Complete the Course (minutes)	27.7	28.1	2.6
Indicated Airspeed (NM/min)	111.0	108.9	8.5
Ground Speed (NM/min)	110.9	108.9	8.5
Radar Altitude (feet)	192.0	184.3	52.5

Table 9. Summary of the Flight Performance Data

The time to complete the course varied little from pilot to pilot, and that corresponds with the small variance in ground speed between the pilots. No wind speed of any kind was entered in to the simulation for any of the pilots, which is why indicated airspeed (IAS) and ground speed (GS) were so close in value.

From the literature review, using a pilot's deviation from the assigned altitude parameters is an acceptable measure of a pilot's performance during a trial. Figures 10 and 11 show the comparison of a pilot's experience level in terms of flight hours and the standard deviation of the altitude during the flight.

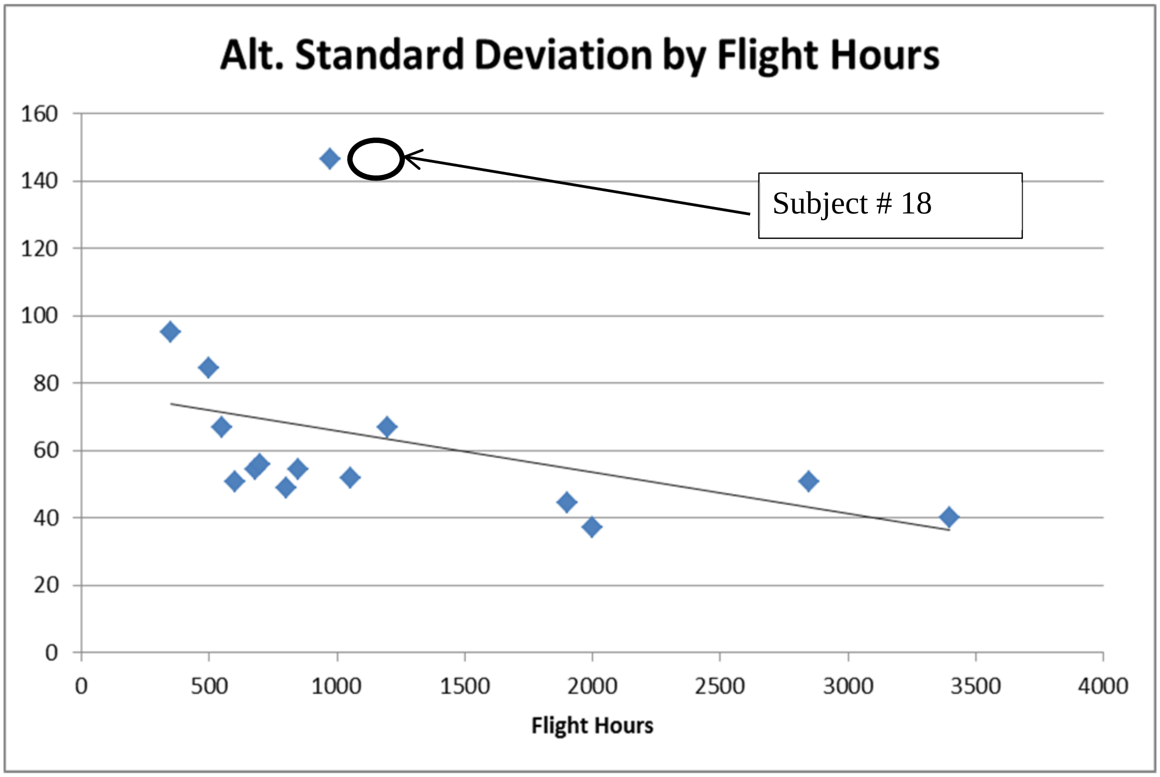


Figure 10. Standard Deviation in Altitude Compared to Flight Hours, *n = 15*

From Figure 10, there appears to be a decreasing trend in the amount of the standard deviation of altitude as pilot experience increases in terms of total flight hours as shown in Figure 16. Subject 18's altitude standard deviation was extreme when compared to those of the other pilots (standard deviation 146.3, 975 flight hours). Figure 11 shows the data with Subject 18's data removed from the data set resulting in a best-fit linear regression that revealed a significant a decreasing trend in altitude variance as pilot experience level increases.

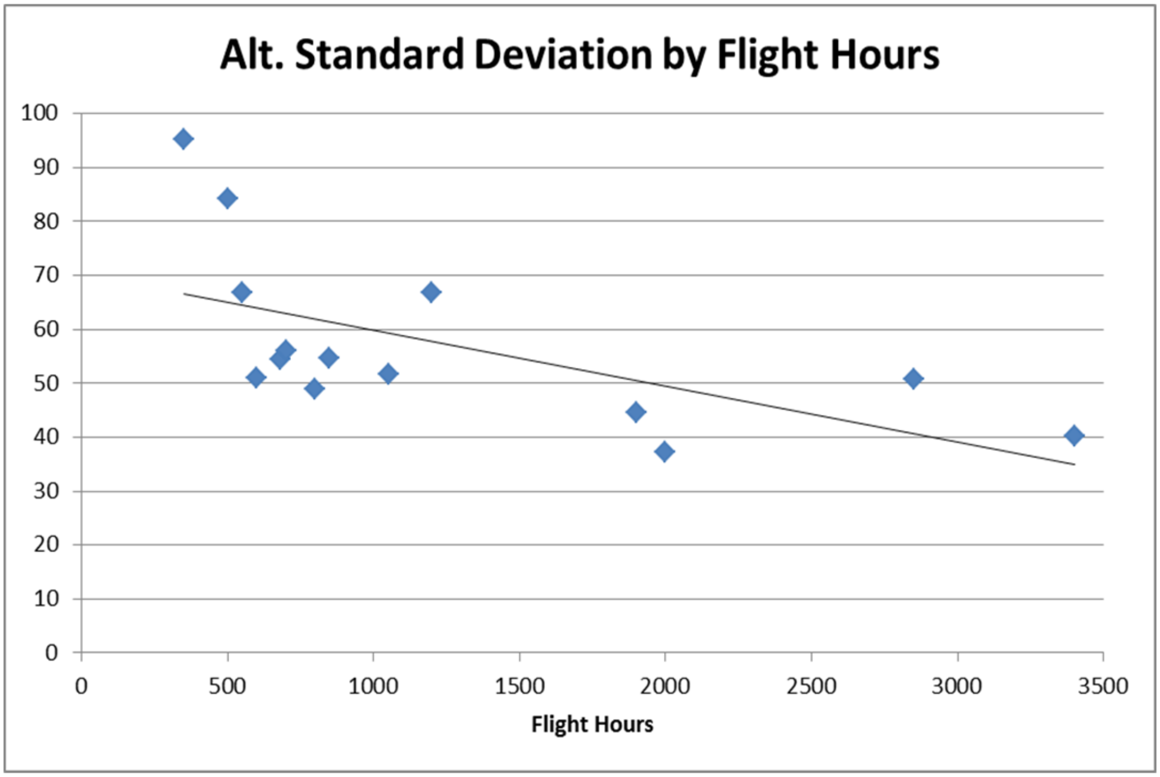


Figure 11. Standard Deviation in Altitude Compared to Flight Hours with Subject 18 Data Removed, $n = 14$

With Subject 18's data removed, the negative association between standard deviation and pilot experience is highly significant (Spearman's *Rho* = - .745, p = .002).

The prior discussion concerning Subject 18 led to another question: was there any correlation between the height at which the pilot flew and the difficulty each pilot had in maintaining the desired altitude above the terrain? If so, one possible explanation is that the higher a pilot flew, the harder it was to maintain a constant altitude above the ground using visual cues. This, when cross-referenced with how often a pilot looked Out-of-the-Window (OTW) using scan data obtained from the FaceLab system, produces a good first glance of how a pilot scans correlates with a metric of pilot performance.

Standard deviation in altitude compared to the average altitude a pilot held throughout the flight is shown in Figures 12 and 13 the next two graphs. Figure 12 includes the data from Subject 18, while Figure 13 removes subject 18's data. This

approach is to ensure that Subject 18's higher than normal altitude variance is not skewing the first look at this performance metric.

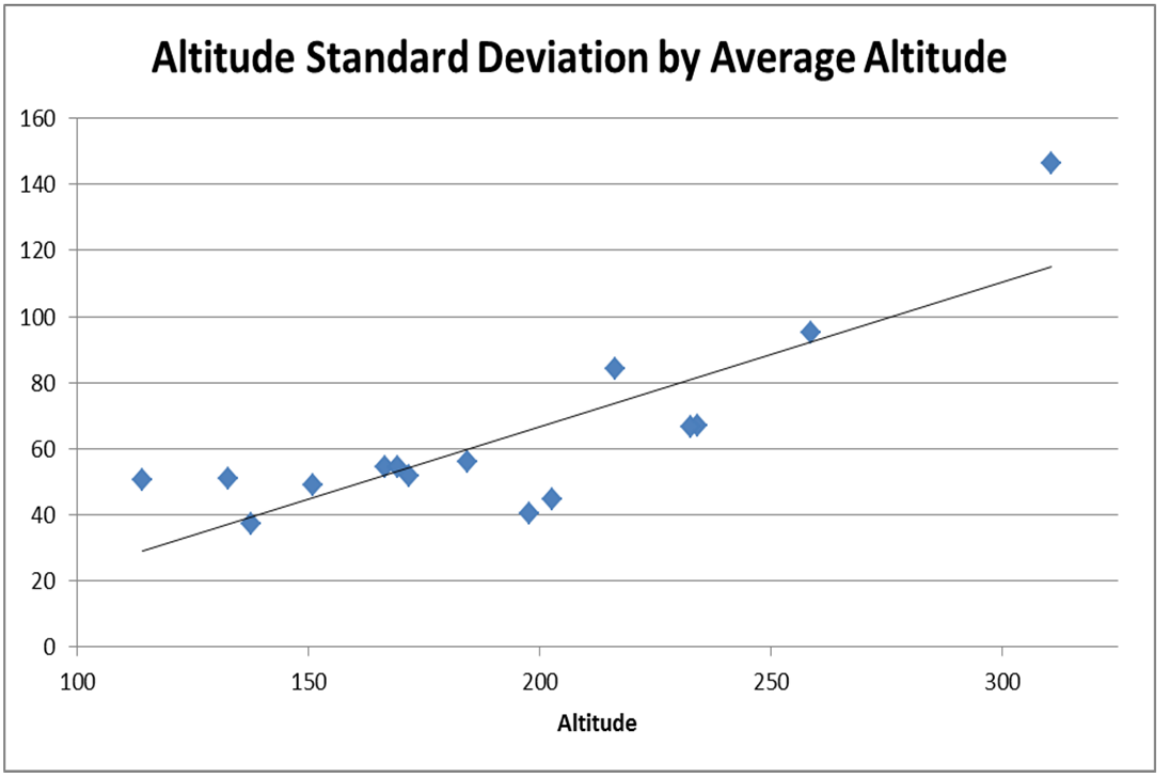


Figure 12. Altitude Standard Deviation by Average Altitude, *n = 15*

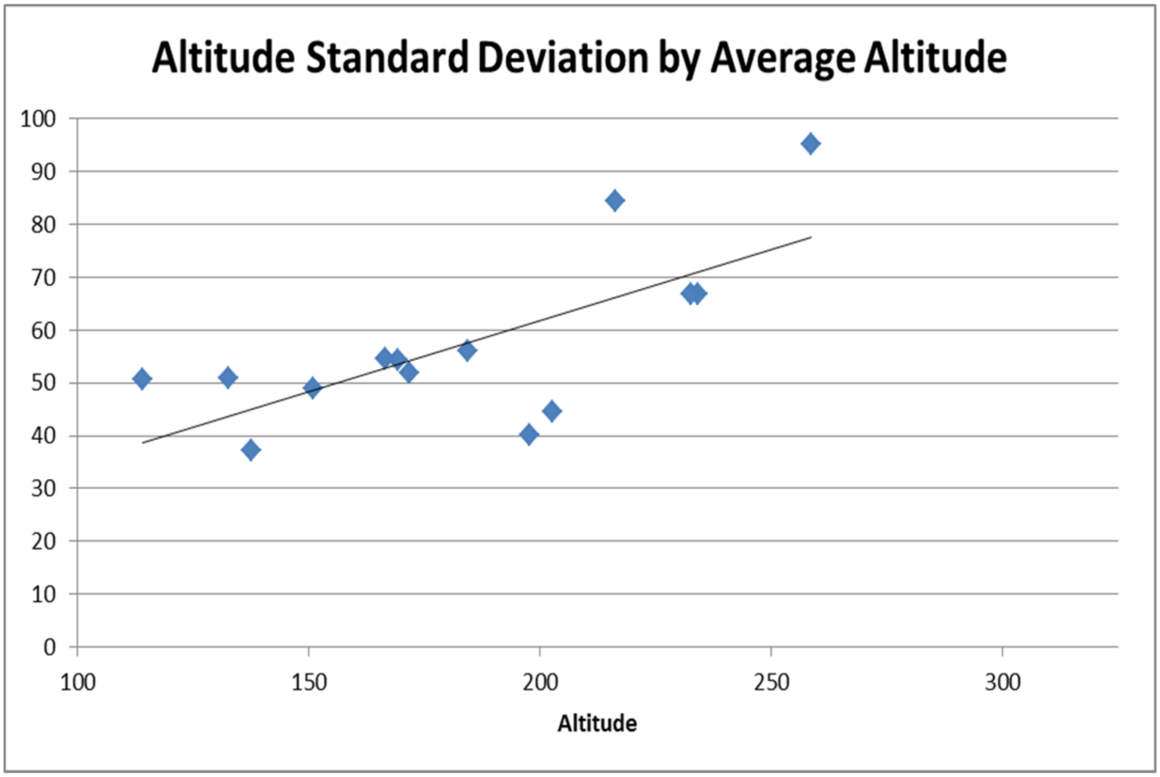


Figure 13. Altitude Variance by Average Altitude with Subject 18's Data Removed, *n =14*

In both cases, the higher a pilot flew, the higher his or her deviation in altitude was.

From this point on, Subject 18's data was retained in the analyses. Spearman's *Rho* changed modestly with the removal of Subject 18's data. It was determined that Subject 18 was not an influential outlier. Therefore, Subject 18's data was included in the analyses for hypothesis testing.

The data suggested two points for further investigation. The pilots with more experience were able to hold a more constant altitude above the ground throughout the flight, and the pilots who flew lower were also able to maintain a more consistent altitude profile. Figure 14 shows a non-significant relationship between the average altitude for the flight and the experience level (in total flight hours) of the pilot flying the route. According to Figure 14, the more experienced pilots flew lower along the route, which was expected.

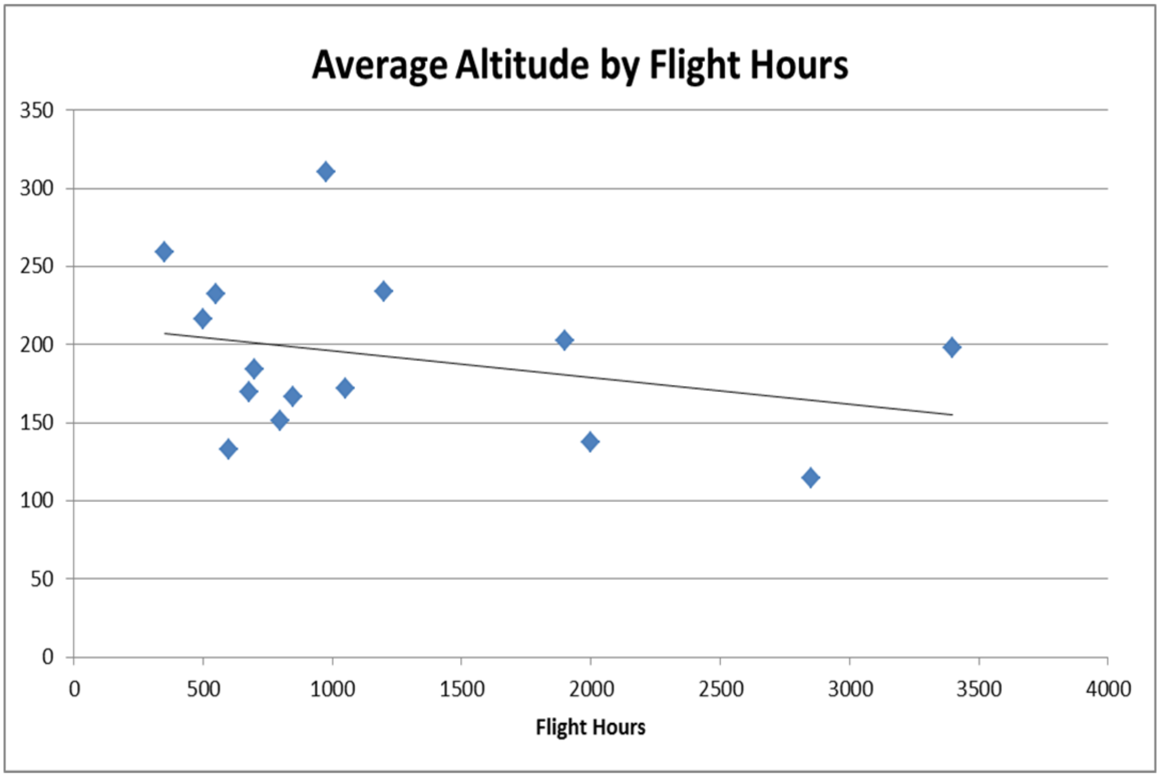


Figure 14. Average Altitude by Flight Hours, *n = 15*

Trend	Spearman's RHO	P-value	Sig. Trend?	Type?
Altitude Std Dev. By Flight Hours	-0.632	0.007	Y	Negative
Altitude Std Dev. By Flight Hours (-18)	-0.745	0.002	Y	Negative
Altitude Std Dev. By Altitude	0.707	0.002	Y	Positive
Altitude Std Dev. By Altitude (-18)	0.639	0.008	Y	Positive
Mean Altitude by Flight Hours	-0.279	0.157	N	Negative

Table 10. Associations between Performance Parameters and Flight Hours

All of the comparisons had significant trends, except for the comparison of altitude and flight hours. For the ease of interpretation, green rows indicate that a significant trend exists ($p < .05$). Blue rows indicate that trend exists, but only within a 10% significance level ($p < .1$), and red rows indicate that a trend was not found within either level of significance ($p > .1$). In this specific case, the green rows indicate that a pilot's performance in terms of altitude deviation improves as his or her experience improves. The red row shows that it was not able to be confirmed, within a 5% level of error, that the more experience pilots flew lower during the flight.

Another and final consideration for the average altitude flown was whether or not the type of experience played a role in how high each pilot flew over the ground during the route. Each pilot was asked to classify the bulk of their flying experience as either “over land” or “maritime” on the demographics survey. Most of the subjects from HSC-21 classified themselves as “over land” pilots. For the last couple of years, HSC-21 has been filling the air-ambulance role in Iraq, a mission that requires all of their flying to be overland. In contrast, the HS-4 pilots classified themselves as “maritime,” since they were fulfilling the classic role of a HS squadron: maritime search and rescue and carrier support. Given these areas of operations for the two squadrons, it was conjectured by the investigator that the pilots that classified the bulk of their experience as “over land” would fly lower along the route since the HSC-21 pilots should be more comfortable over land than the HS-4 pilots. This conjecture was not supported by the data collected from TOFT-2’s de-brief system, $t\,(16) = .638$, $p = .547$. The pilots who classified themselves as “over land” pilots (7) flew at an average height above ground of 201.3 feet (sd 56.1), while the “maritime” (8) pilots flew at an average height above ground of 183.9 feet (sd 39.0). This somewhat surprising result could be followed up in a study that had larger numbers of pilots in each group.

2. Eye Tracking Measures Analysis

Given that altitude standard deviation has been established as a performance measure, analysis was done to see how the pilot’s performance related to scan rate and dwell times. Since the scan rates and the average dwell times were computed separately, a check was needed to make sure the two had an inverse relationship. That is, as the scan rate increases for each pilot, their average dwell time should decrease. This correlation is only logical since an increased scan rate indicated that the pilot was spending less time on each object, and moving between them more frequently. The following graph displays this inverse relationship.

	Scan Rate (shifts/sec)	Dwell Time (sec)
Mean	2.30	0.82
SD	0.96	0.45
Median	2.09	0.76

Table 11. Summary of the Scan Rate and Dwell Time Data

The dwell time, or the average time a pilot spent looking in any one particular area, was measured in seconds. Scan was measured in shifts per second, as it is the number of times a pilot shifts his scan from one place to another during a measured time unit.

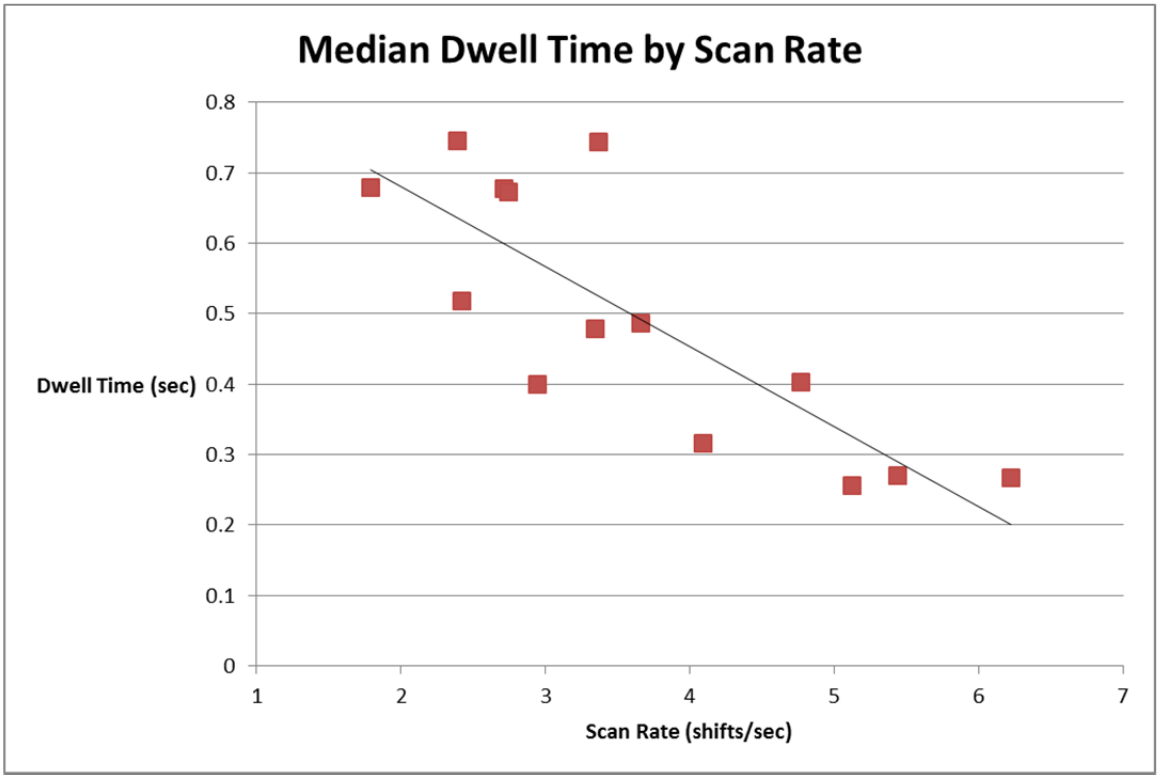


Figure 15. Median Dwell Time by Scan Rate, $n = 14$

Dwell times and scan rates had an inverse relationship. The next step was to determine whether or not there were trends in dwell times and scan rates as they relate to

pilot experience level. Figures 16 and 17 show this. In both cases, there was a good deal of variability for both scan rate and fixation time for those pilots with less than 1500 hours.

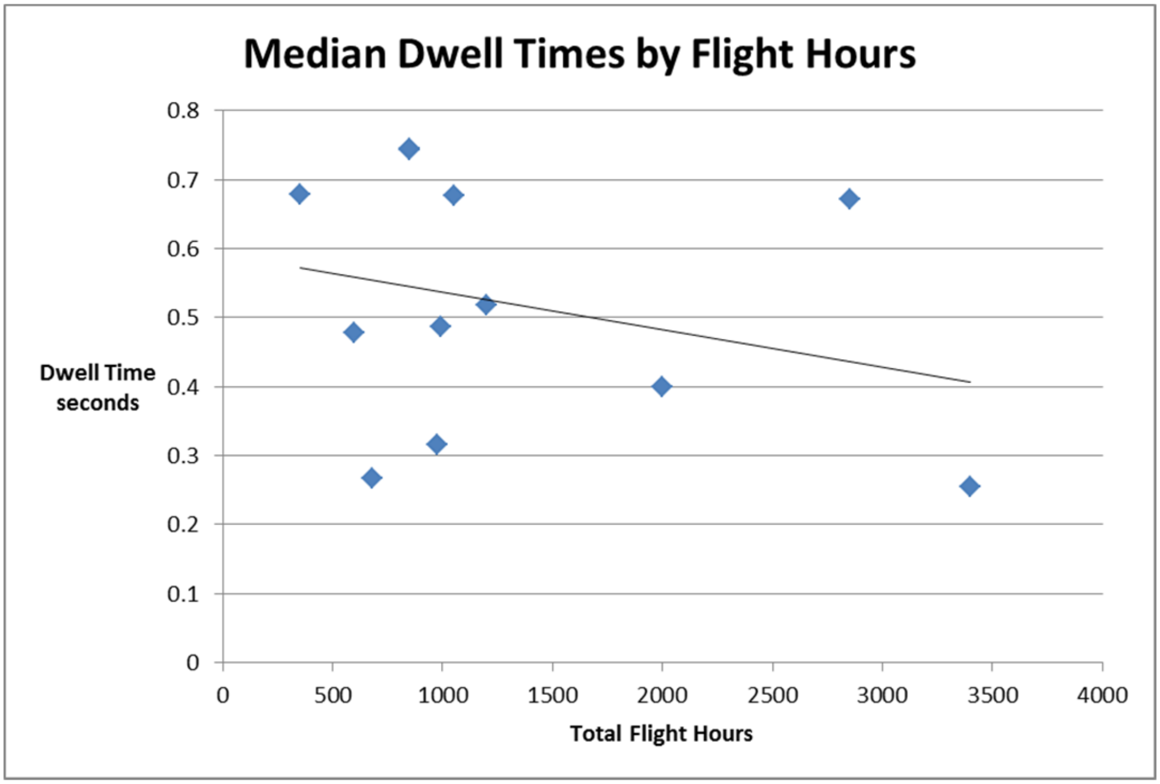


Figure 16. Median Dwell Time by Flight Hours, *n = 14*

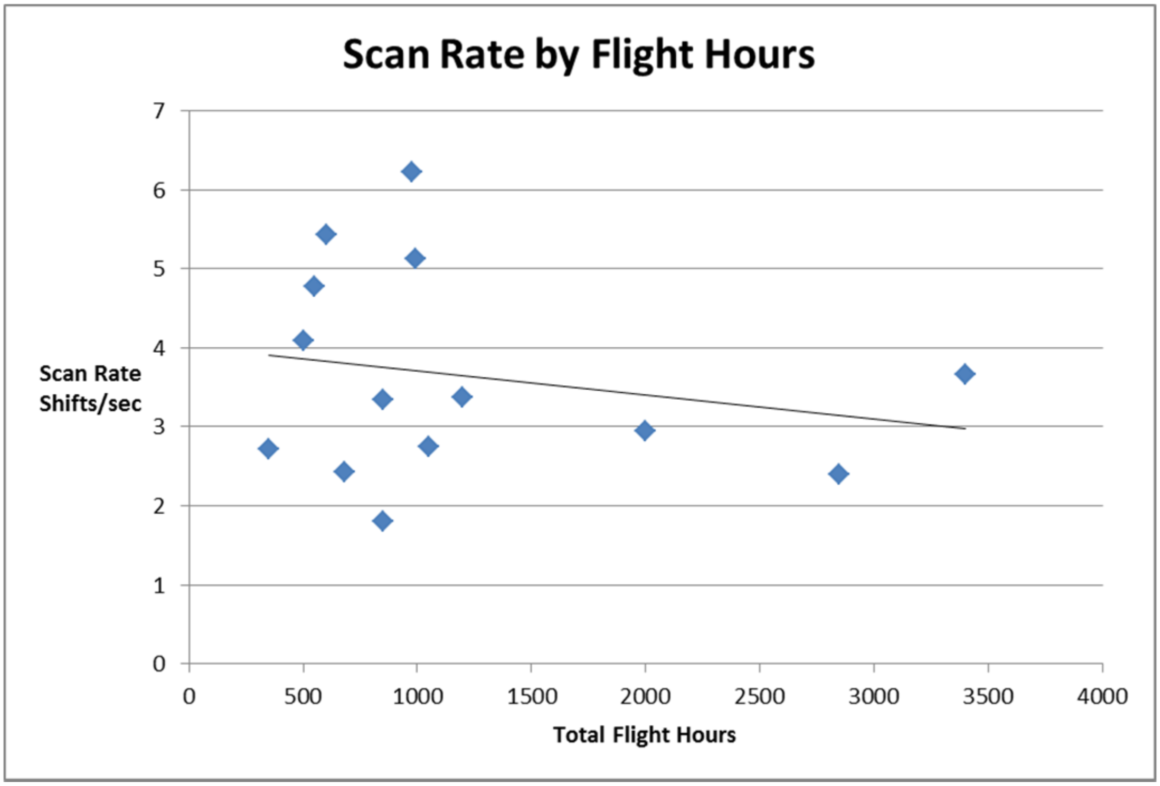


Figure 17. Scan Rate by Flight Hours, *n = 14*

Since the dwell times and scan rates had an inverse relationship, it follows that this relationship would carry over into the comparison of the two with a pilot's flight hours. The interesting discovery was a positive but not significant association shown in table 12 (*Rho* = .231, *p* = .236) between dwell time and experience, which contradicted the regression analysis shown in Figure 16, along with a non-significant negative association (*Rho* = .229, *p* = .216) between scan rate and experience. Neither scan rate nor fixation time predicted variability in altitude with a 5% significance level, although trends were present when the data was presented graphically in Figures 18 and 19.

Trend	Spearman's RHO	P-value	Sig. Trend?	Type?
Median Dwell Times by Flight Hours	0.231	0.236	N	Positive
Scan Rates by Flight Hours	-0.229	0.216	N	Negative

Table 12. Association between Eye Tracking Parameters and Flight Hours

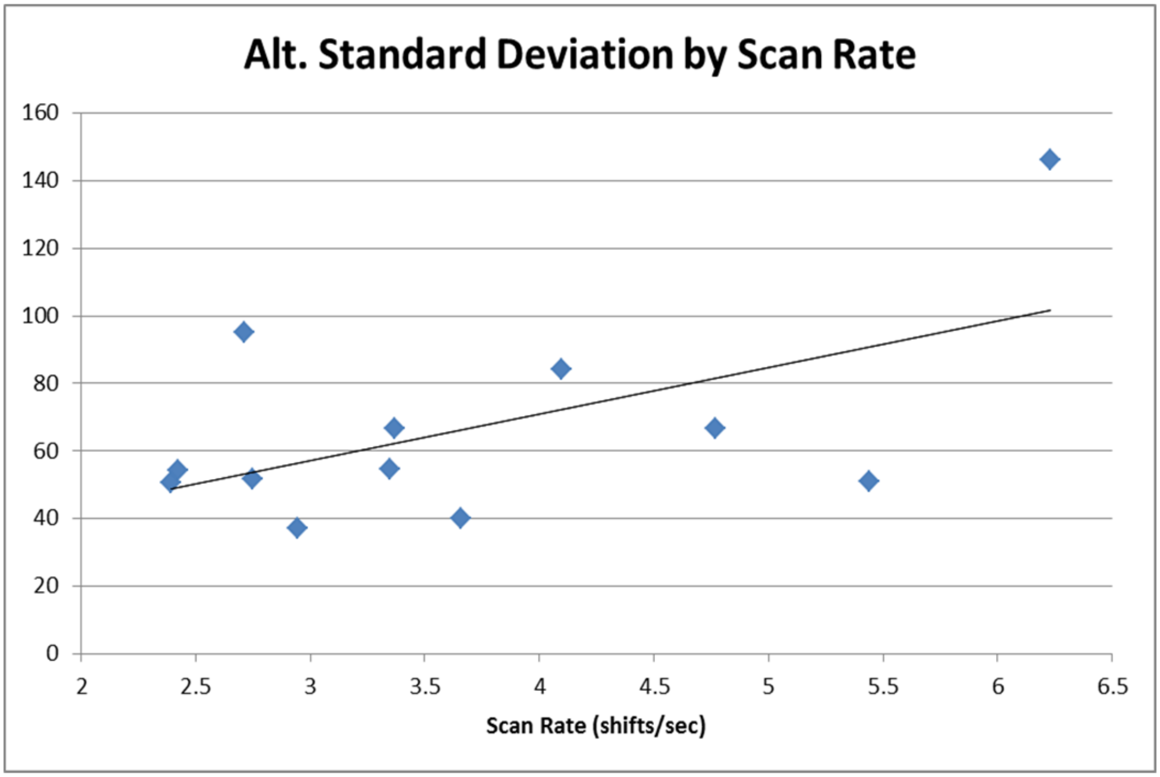


Figure 18. Altitude Standard Deviation by Scan Rate, *n = 12*

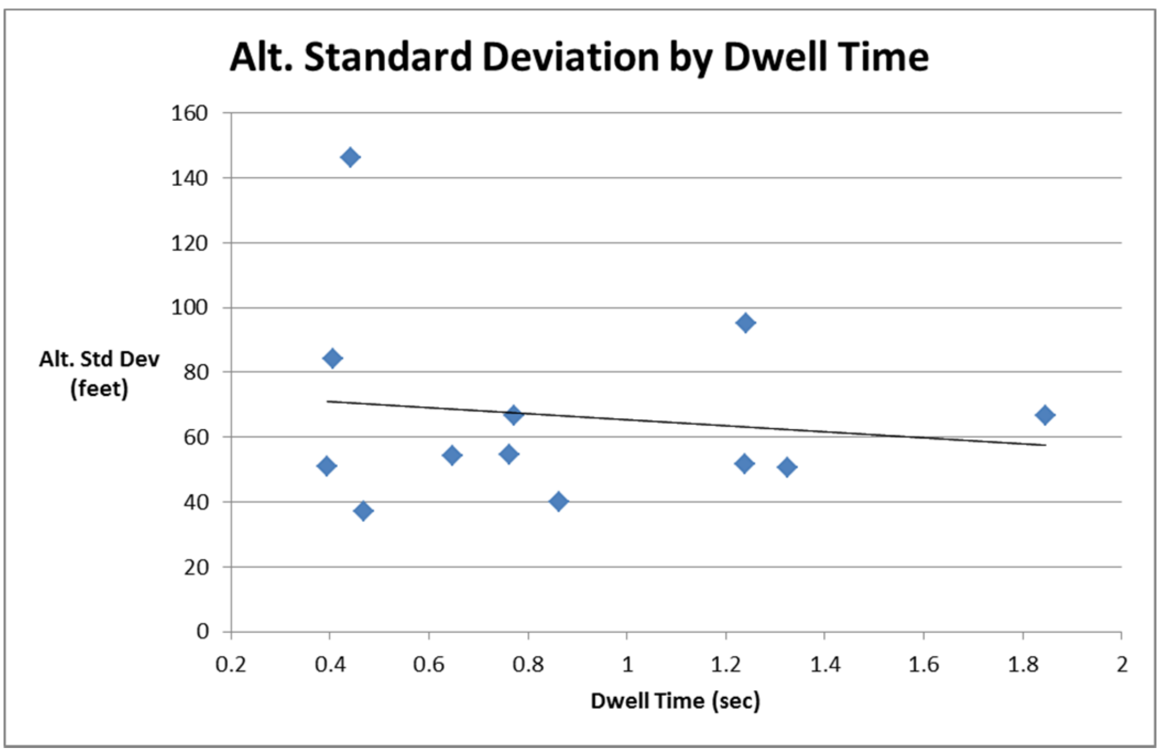


Figure 19. Altitude Standard Deviation by Fixation, *n = 12*

Trend	Spearman's RHO	P-value	Sig. Trend?	Type?
Altitude Std Dev by Scan Rate	0.133	0.302	N	Positive
Altitude Std Dev by Median Dwell Time	-0.315	0.160	N	Negative

Table 13. Association between Eye Tracking Parameters and Altitude Standard Deviation

Another metric that was obtained from the FaceLab data was dwell time, the amount of time a pilot spends looking at one object. The cumulative dwell times delineate how much time a plot spends looking at a particular object during the duration to the flight, such as the instrument panel (Inst_Disp) or the aircraft status panel (Map). In the data preparation section, this data was presented as the percentage of the total time of the flight the pilot spent looking at a particular object. In Table 14 and Figure 20, these percentages were compared to pilot experience to gain some insight as to where the more experienced pilots spend most of their time looking at.

	Inst_Disp (%)	MAP (%)	OTW (%)
Mean	58.43	7.40	34.17
SD	19.57	17.52	19.68
Median	58.26	0.26	39.00

Table 14. Summary of the Scan Location Parameters

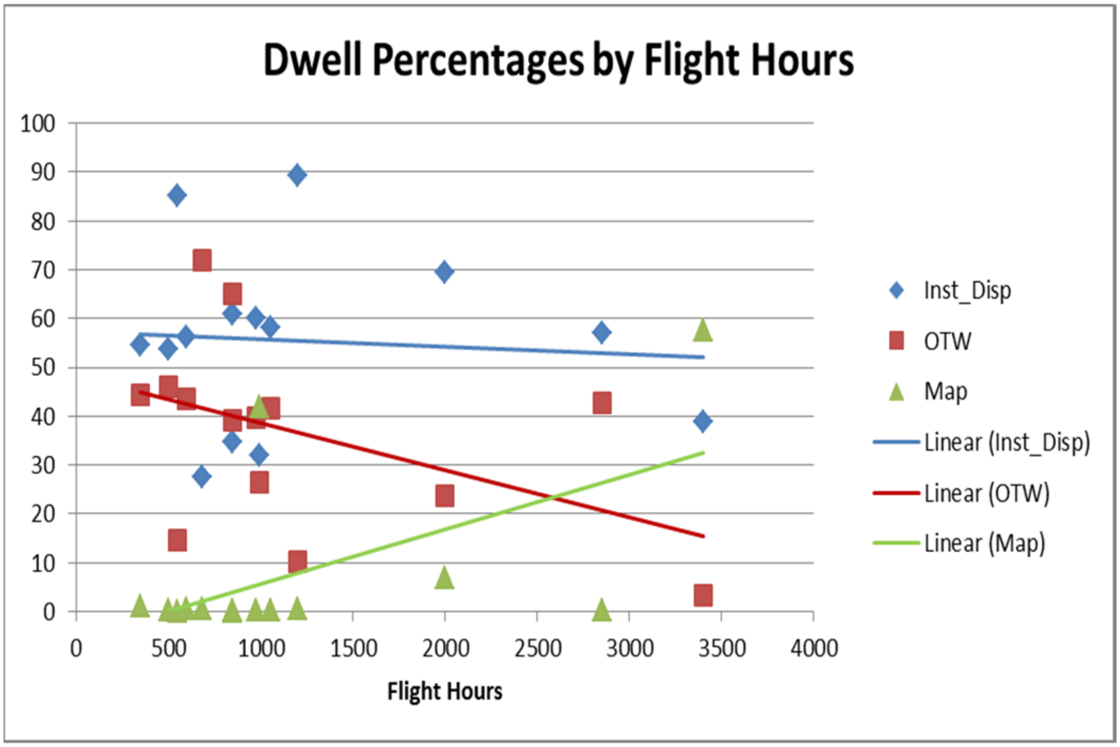


Figure 20. Dwell Percentages by Flight Hours, *n = 14*

The results in Figure 20 show that the amount of time a pilot spends looking at the instrument display was consistent across experience levels. However, the more experienced pilots spent far less time looking out of the window, or flying with a visual scan, than the less experienced pilots. These more experienced pilots seemed to swap the amount of time spent looking outside with time spent looking at the aircraft diagnostics screen (called the Map display in FaceLab). Subject 8, the most experienced pilot of the subject group (3400 hours total flight time) was the most extreme case, with a significant amount of time (57.7%) spent scanning the aircraft diagnostics (AD) display. The next subject in decreasing order according to AD display dwell time was Subject 16 (41.8%, 991 hours total flight time). All other subjects spent less than 7% of the time scanning the AD display.

To see what kind of effect these two extreme cases had, they were removed in the generation of Figure 21. The effect of the dwell percentage on the AD (MAP) display is nearly removed, leaving the trends in the amount of time a pilot scans the instrument display vs. out of the window as experience increases more evident. Table 15 contains

trend analyses from both cases, with Subjects 8 and 16 and without. The trend for the amount of time pilots spend looking at the AD (Map) display is almost non-existent without Subjects 8 and 16. Other trends, such as dwell percentages for the instrument display and out-of-the-window, become more evident (in Table 15) with their data removed.

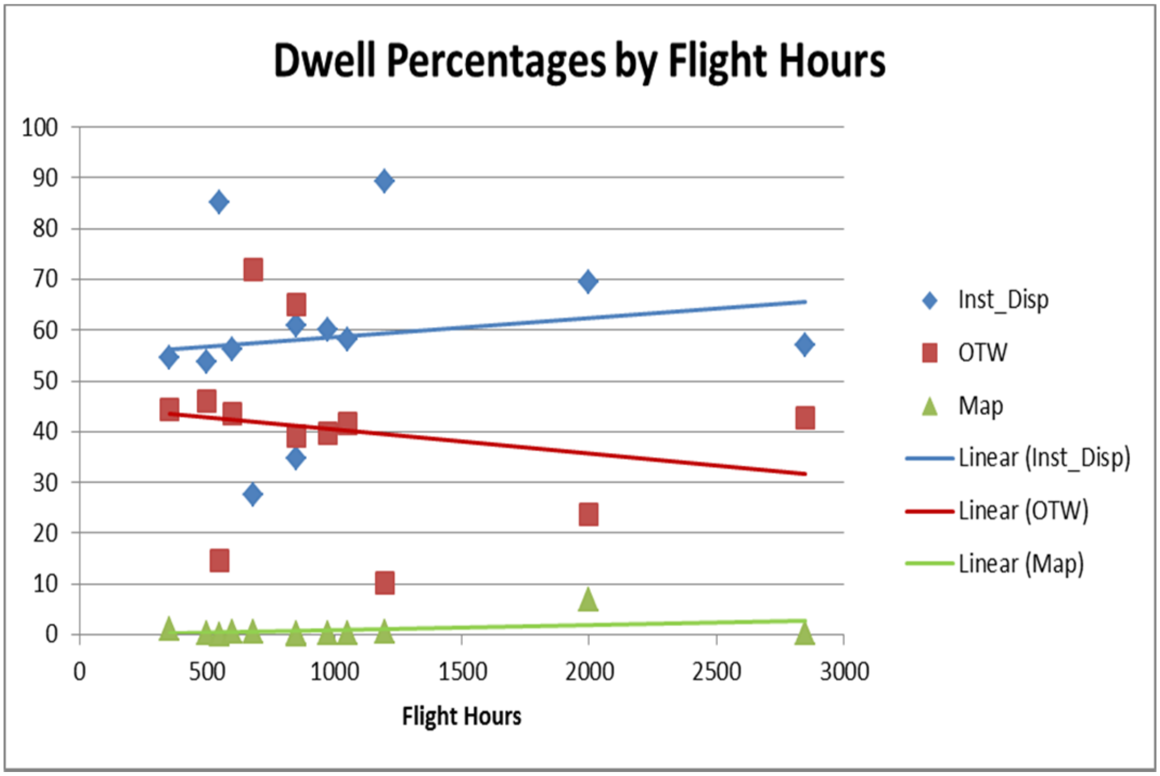


Figure 21. Dwell Percentages by Flight Hours with Subjects 8 and 16 Data Removed, *n =12*

Trend	Spearman's RHO	P-value	Sig. Trend?	Type?
Dwell %'s by Flight Hours (OTW)	-0.563	0.018	Y	Negative
Dwell %'s by Flight Hours (OTW, - 8, 16)	-0.430	0.080	N	Negative
Dwell %'s by Flight Hours (Inst_Disp)	0.165	0.287	N	Positive
Dwell %'s by Flight Hours (ID, w/o -8,16)	0.431	0.081	N	Positive
Dwell %'s by Flight Hours (Map)	0.275	0.171	N	Positive
Dwell %'s by Flight Hours (Map, w/o -8,16)	-0.007	0.501	N	Negative
Dwell % by Altiude Deviation (OTW)	0.401	0.096	N	Positive
Dwell % by Altiude Deviation (Inst_Disp)	-0.301	0.165	N	Negative

Table 15. Association between Eye Tracking Parameters and Flight Hours, Altitude Deviation

According to Table 15, the more experienced pilots spent less time looking out of the window than the less experienced pilots. This was the only trend that was verified within a 5% significance level. Of note were the comparison of the OTW dwell percentages by flight hours, the ID dwell percentage by flight hours, and the dwell percentage on OTW by altitude standard deviation, all of which showed trends within the 10% significance level (shown in Figure 22).

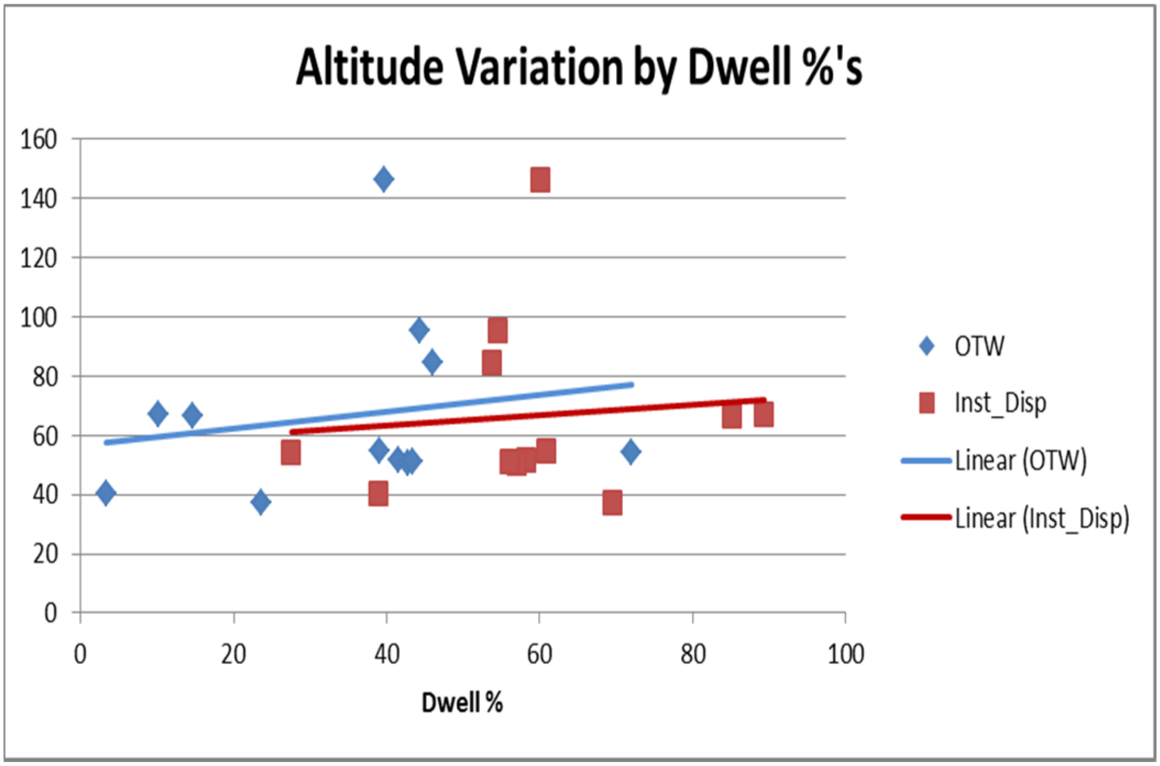


Figure 22. Altitude Deviation Compared by % of Dwell Time, *n =12*

In both preliminary looks at the dwell percentage data, the trend for more experienced pilots to look out of the window less remained consistent. When data from Subjects 8 and 16 were removed, along with their tendency to scan the AD display more than any other pilot in the study, the decrease in the amount of time spent scanning visually was traded for more time spent scanning the instrument panel. No similarities

existed between Subjects 8 and 16 other than this tendency. Why these two pilots scanned the AD display significantly more than any other of the pilots in the study is beyond the scope of this study.

a. Fixation Analysis

Fixation, as an event, occurs when a pilot keeps her scan on an area of interest for more than 70 milliseconds. Using Microsoft Excel to further analyze the world-view data, each pilot's data was processed to find events that fit this criteria. FaceLab had a frame rate of 60 frames per second. In 70 milliseconds, 4.2 frames passed. Any difference that was greater than 4.2 frames was counted. That is, any time a change from one area to another (an event recorded in the previous analysis to determine scan rate) took more than 4.2 frames, the event that the shift was *from* was counted as the fixation area of interest. Table 16 shows a summary of the fixation data collection.

	Fixation			
	IP	MAP	OTW	No Fix
Mean	629.80	83.67	1094.47	4518.67
SD	444.08	163.48	817.76	1247.64
Median	669.00	15.00	963.00	4375.00

Table 16. Summary of Fixation Events

Table 16 shows a summary of the fixation events recorded using the 70 millisecond criteria. Each area is represented; "IP for instrument panel, "MAP" for the aircraft diagnostic display, and "OTW" for scanning out of the window. "No Fix" represents the number of recorded events that did not meet the fixation criteria. These are the scan events that were shorter than 70 milliseconds. Figure 23 further summarizes the data in a box plot.

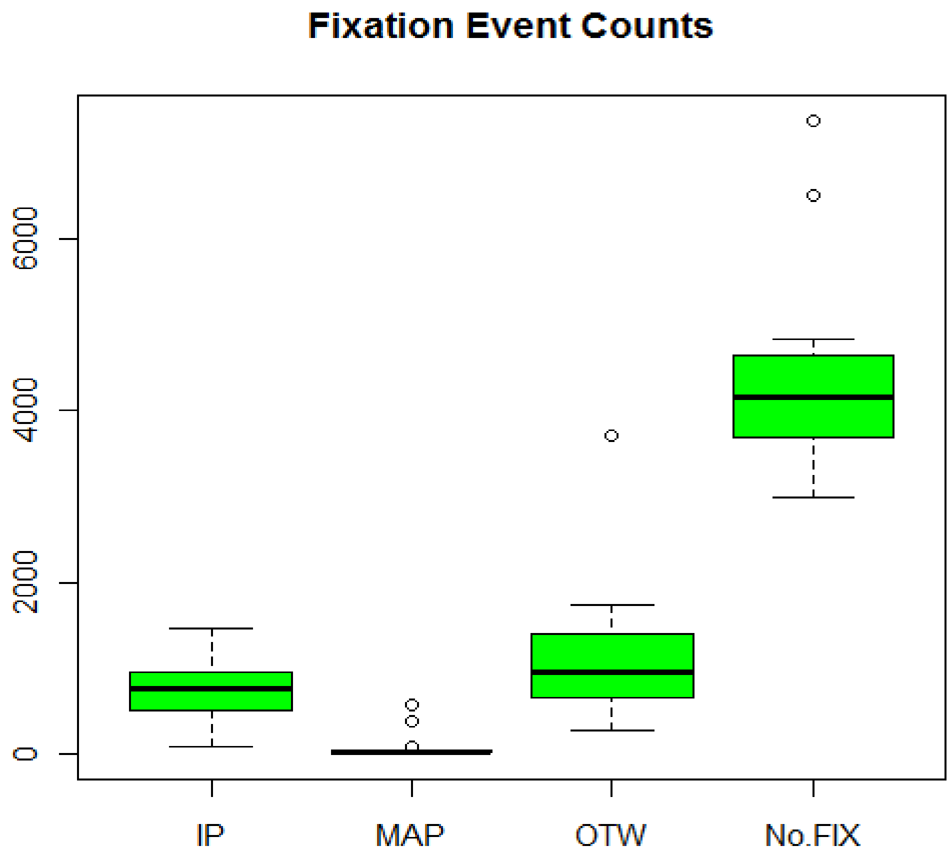


Figure 23. Boxplot of Fixation Events for All Pilots

To further explore whether or not pilots scan techniques differ as their experience increases, the number of fixation events for each pilot were compared on a scale of flight experience via the amount of flight time accrued. Figure 23 shows the number of fixation events recorded using the 70 millisecond criteria for each area of interest. As flight experience increases, there is a slight increase in the number of fixations on each area of interest. The amount of fixations on the instrument panel showed the most significant increase with flight experience, with a slope of 0.19. Out of the window was 0.11, and the Map display's slope was 0.05.

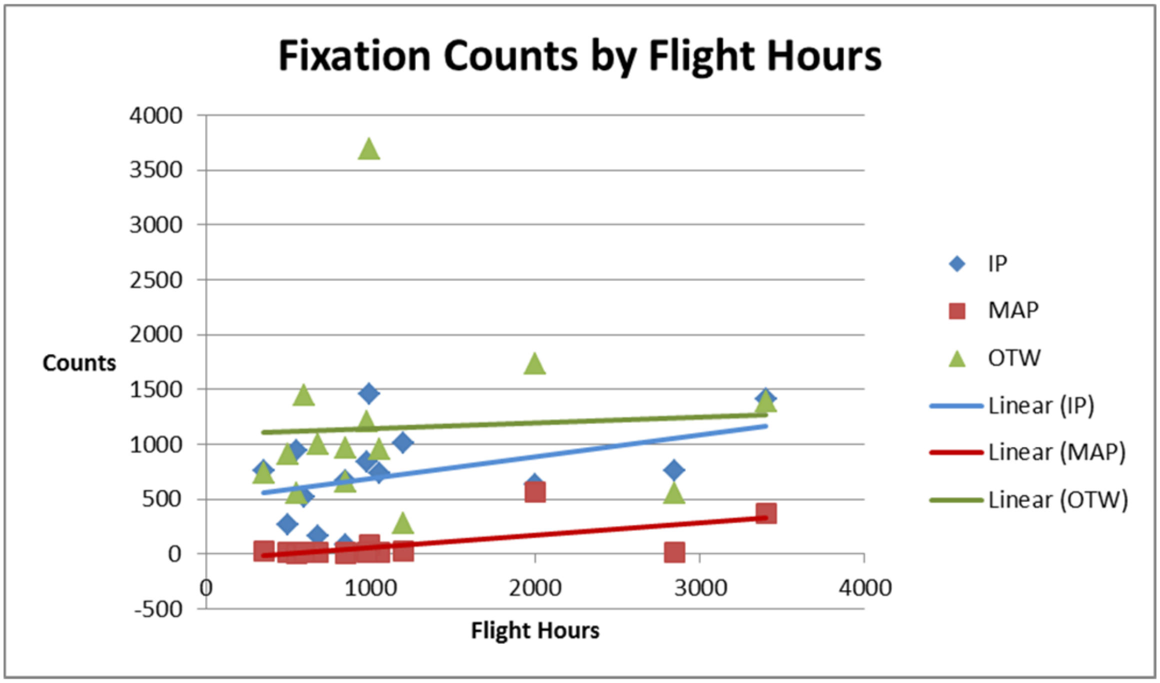


Figure 24. Graphs of the Fixation Counts by Flight Hours

The increase in the number of fixations in each are of interest seemed incomplete without including the "no fixation" event. These are accounted for in Figure 25. As the number of fixations increased, the number of non-fixation events decreased. This relationship implies that pilots with slower scan rates tended to fixate more, spending more time scanning each area of interest for information. This is further confirmed in Figure 26 and Table 18.

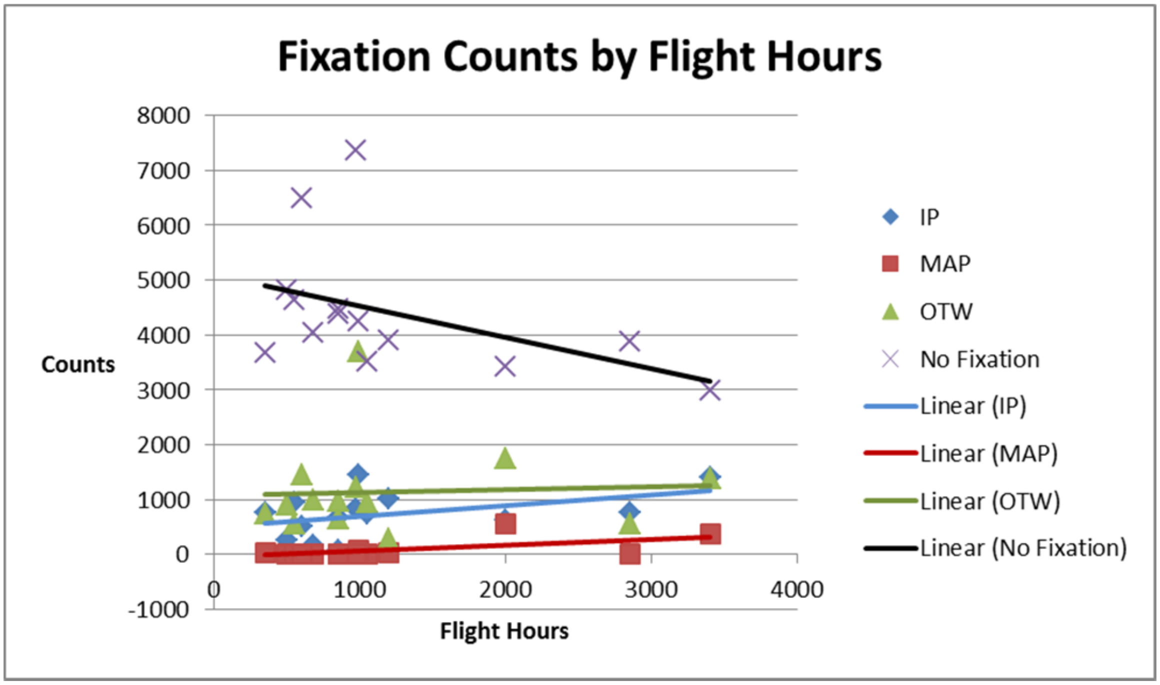


Figure 25. Fixation Counts by Flight Hours, with the "No Fixation" Event Included

Figure 25 shows the addition of the "no fixation" events. The linear fit had a slope of -0.57, showing a significant decrease when compared to pilots' experience in terms of flight hours. The results from further analysis using Spearman's *Rho* to detect trends is shown in Table 17.

Trend	Spearman's RHO	P-value	Sig. Trend?	Type?
IP fixation events by Flight Hours	0.405	0.075	N	Positive
MAP fixation events by Flight Hours	0.276	0.169	N	Positive
OTW fixation events by Flight Hours	0.187	0.261	N	Positive
"No fixation" events by Flight Hours	-0.585	0.014	Y	Negative

Table 17. Fixation Event Analysis Results

"No fixation" events decrease significantly with an increase in experience based on flight hours. The earlier analysis on the scan rate did not produce a significant trend when compared to the number of flight hours a pilot had. Yet, using the no fixation events as a metric to determine how fast a pilot was looking from one area of interest to another, there seems to be a correlation.

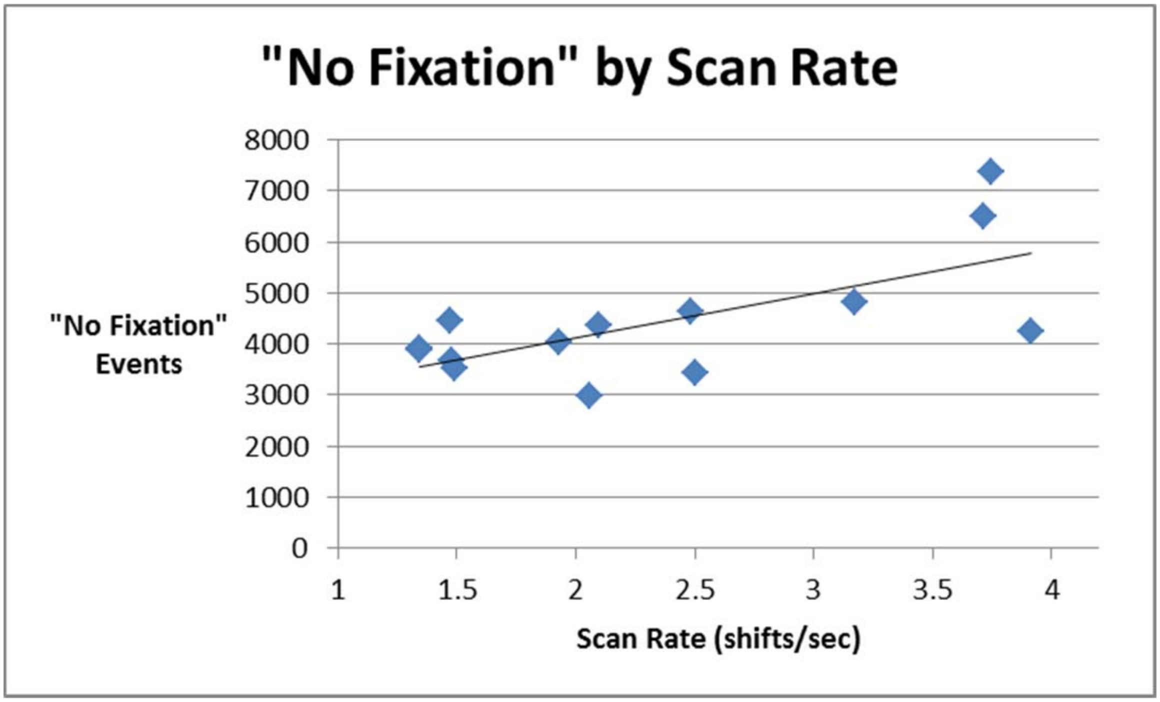


Figure 26. No Fixation by Scan Rate Graph

Figure 26 shows the correlation between scan rate and the number of “no fixation” events recorded. As expected per the definitions, the faster a pilot’s scan rate, the less time he spent looking at any one object. Hence the increased number of “no fixation” events. Table 12 did not show a significant relationship between scan rate and flight hours. Further research is needed to determine if the “no fixation” events can be used as a metric to measure a pilot’s scan rate.

Trend	Spearman's RHO	P-value	Sig. Trend?	Type?
"No Fixation" by Scan Rate	0.494	0.037	Y	Positive

Table 18. “No Fixation” by Scan Rate Trend Analysis

The last remaining item from the data preparation to receive an initial look was the scan direction data that was created from the FaceLab collection efforts. This analysis was done to gain insight in to a pilot’s scan direction. By looking at the number of transitions that occur from one area to another in Table 19 and Figure 27, a general idea of scan direction could be ascertained. The scan shift event codes were included in Table 20 for ease of interpretation.

	IM	IO	MI	MO	OI	OM
Mean	103.13	1823.53	102.87	15.33	1823.93	15.07
SD	106.97	850.87	105.38	25.27	852.50	23.56
Median	62.00	1571.00	62.00	6.00	1571.00	9.00

Table 19. Summary of Scan Shift Even Counts

Scan		
Code	From	To
IM	Inst_Disp	MAP
IO	Inst_Disp	OTW
MI	MAP	Inst_Disp
MO	MAP	OTW
OM	OTW	MAP

Table 20. Scan Transition Codes

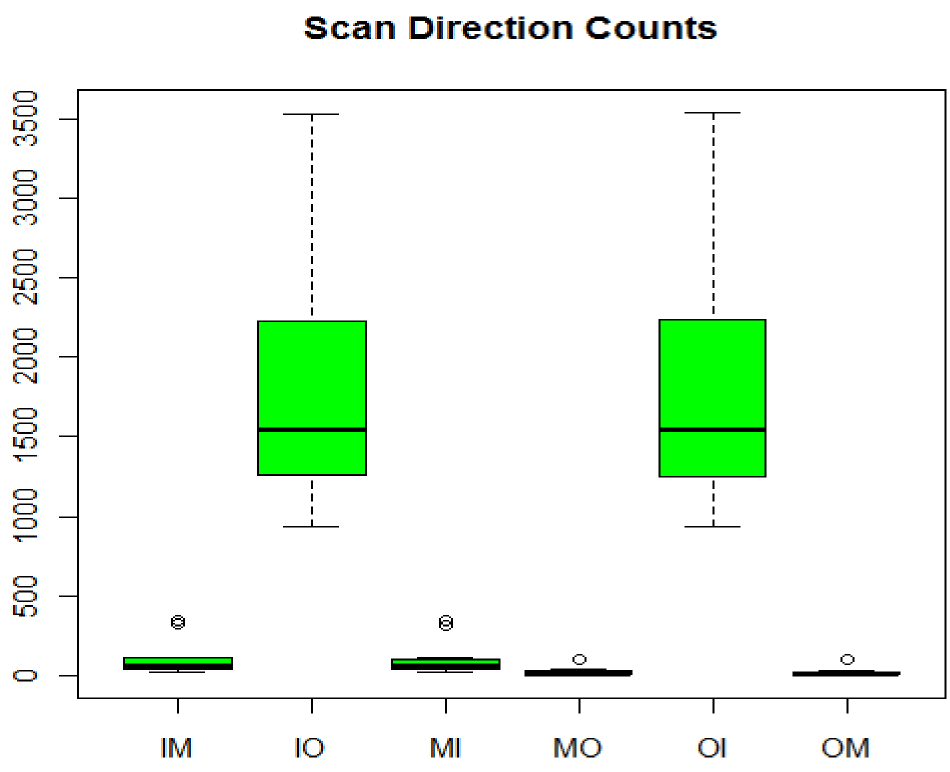


Figure 27. Boxplots of the Scan Shift Directions, $n = 14$

Comparison	T	P-value	Difference?
Scan Direction counts			
IO vs MO	8.742	0.000	>
OI vs OM	8.694	0.000	>
IM vs OM	3.41	0.002	>
MI vs MO	3.42	0.002	>

Table 21. Scan Direction Comparisons Using a Paired T-Test

From the data in Table 21 and Figure 27, a simple diagram was constructed in Figure 28 in order to show where the bulk of the scan activity for all of the pilots existed. Pilots were, for the most part, scanning in between their instrument displays and out of the window. There were more transitions from the instrument display to the AD display (in either direction) than from the AD display to out of the window. It seemed that pilots spend most of the flight scanning out of the window and the instrument display, with a few looks at the aircraft diagnostic page from the instrument panel.

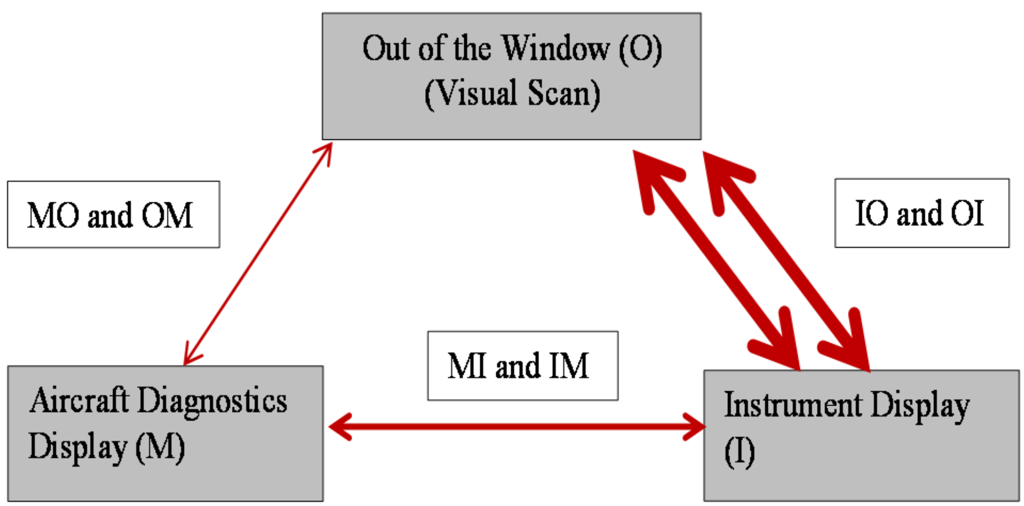


Figure 28. Diagram to Depict the Intensity of Scan Activity between Objects. The Width and Number of the Arrows Indicates the Frequency of the Scan Shift Event

D. SUMMARY OF RESULTS

Analysis of the results obtained from the eye tracking system indicates a decreasing relationship between scan rate and pilot experience, indicating that the scan rate decreases as a pilot becomes more experienced. The analysis uses altitude variance as a measure of performance. Results indicate that higher scan rates correlate with higher degrees of variance in the altitude, indicating that a quicker scan does not necessarily result in better performance. The higher experienced pilots show a lower altitude variance overall (they were more consistent in maintaining a constant altitude above the ground), yet those pilots all exhibited slower scan rates.

E. HYPOTHESIS TESTING

1) H0: There is no association between eye scan pattern (number of fixations, dwell durations, percent time looking out the window (OTW), scan rate between OTW and instrument panel) and total flight hours.

HA: Eye scan patterns will be more efficient for pilots who have higher total flight hours (greater number of fixations, shorter dwell durations, less time looking OTW, and faster scan rate between OTW and instrument panel).

Result: Fail to reject the null hypothesis. Although Spearman's *Rho* indicates trends between eye scan pattern and total flight hours and the dwell times ($r_{spearman}$ = .310, p = .14) and scan rate ($r_{spearman}$ =-.229, p =.216), none of the results were significant. The dwell percentage for out of the window scanning ($r_{spearman}$ = -.563, p = .018) did show a significant trend with an increase in flight experience, however it was the only scan region that demonstrated such a trend (Table 15).

2) H0: Eye scan parameters (number of fixations, dwell duration, percent time OTW, scan rate between OTW and instrument panel) will not predict a pilot's ability to accurately maintain assigned flight parameters during a navigation event.

HA: Eye scan parameters (number of fixations, dwell duration, percent time OTW, scan rate between OTW and instrument panel) are a reliable predictor of a pilot's ability to accurately maintain assigned flight parameters during a navigation event.

Result: Fail to reject the null hypothesis. Spearman's *Rho* in this case only showed very slight trends when either dwell times ($r_{spearman}$ =-.021, p = .302),

scan rates ($r_{spearman}$ =.133, p = .478) or were compared with the altitude standard deviation. From Table 15, only the amount of time a pilot spends looking out of the window had any significance when compared to altitude deviation ($r_{spearman}$ =-.401, p = .096), but the amount of time a pilot spent looking at the instrument display did not change the amount a pilot deviated from the altitude with any significance ($r_{spearman}$ =-.301, p = .165).

3) H0: Eye scan parameters (number of fixations, dwell duration, percent time OTW, scan rate between OTW and instrument panel) will not predict occurrence of CFIT.

HA: Eye scan parameters (number of fixations, dwell duration, percent time OTW, scan rate between OTW and instrument panel) will predict occurrence of CFIT.

Result: There were no occurrences of CFIT during any of the trials. Therefore, this hypothesis was not tested.

F. EXPLORATORY ANALYSIS

To explore some additional questions that arose from the preliminary results, exploratory analyses were conducted on two pilots who stood out from the others in terms of their performance or eye scan pattern. Additionally, because the older pilots tended to have more TFH, I also explored whether pilot age was associated with flight performance or eye scan pattern.

1. Subject 18

This pilot’s average air speed (116.9 KIAS) and ground speed (116.9 KIAS) were both within one standard deviation of the average air speed and ground speed of all of the pilots (109.80 KIAS, sd = 7.79 for both air and ground speed). However, Subject 18’s average radar altitude (310.6 feet, sd = 146.34), the main measure of flight performance was much higher than the average radar altitude for all of the pilots (195.0 feet, sd = 57.64). At this higher altitude, it would have been difficult to pick out prominent land features that were more easily discernable at lower altitudes. Flying at a higher altitude during these flights was an issue, since the pilots were asked to simulate a tactical operating environment. That is, the lower altitude may have been necessary during this particular mission to increase the chance of survival. Subject 18's eye scan data was

examined to see if there was any explanation for Subject 18's higher than average altitude during the flight. Subject 18 spent 60.1% of the flight scanning the instrument display compared to 58.4% (sd=19.6%) for all of the pilots. Airspeed control was the only other measure of pilot performance recorded during the trials. Subject 18 performed within the standard deviations for all of the pilots in this metric. From Table 10 and Figure 12, the only other possible explanation for the increased deviation is the high altitude Subject 18 held throughout the flight.

2. Subject 8

Subject 8 reported a very low over land hour amount, just 100 hours, compared to the mean of all the pilots' reported over land flight hours of 636.67 (sd = 386.24). Subject 8 had 3400 total flight hours (mean of all pilots = 1273.18, sd = 881.01), and was by far the most experienced pilot of the group.

	Scan Rate	Fixation	Inst_Disp	MAP	OTW
Subj 8	2.06	0.86	38.96	57.66	3.37
Mean	2.30	0.82	58.43	7.40	34.17
SD	0.96	0.45	19.57	17.52	19.68

Table 22. Performance Data for Subject 8

Table 22 shows that Subject 8 performed within a standard deviation with regards to scan rate, fixation, and the percent of time spent scanning the instrument display. Subject 8 spent a great deal more time scanning the map, and less time scanning out of the window than the other pilots. The "MAP" was actually the aircraft diagnostics page. No emergency situations were presented during the flight that would have caused Subject 8 to devote more time scanning the diagnostics page than any other pilot. Also, the small amount of time she spent scanning out of the window is alarming. This indicates that Subject 8 flew the route primarily by referencing the flight instruments.

3. Pilot Age

From Figures 16 and 17, and Table 12 (Scan Rate and Fixation by Flight Hours), it appeared that pilots actually slow down their scan as they accrue more experience. There were a few possibilities for this initial discovery. One of the metrics available was each pilot's age, which was given by each pilot when he or she filled out the pilot demographics survey. A pilot should accrue more flight time as they get older simply because they've had more years to fly (shown in Figure 29). Spearman's *Rho* indicated that there was not a significant (with 5%) association between total flight hours and pilot age (Spearman's = .430, p = .063).

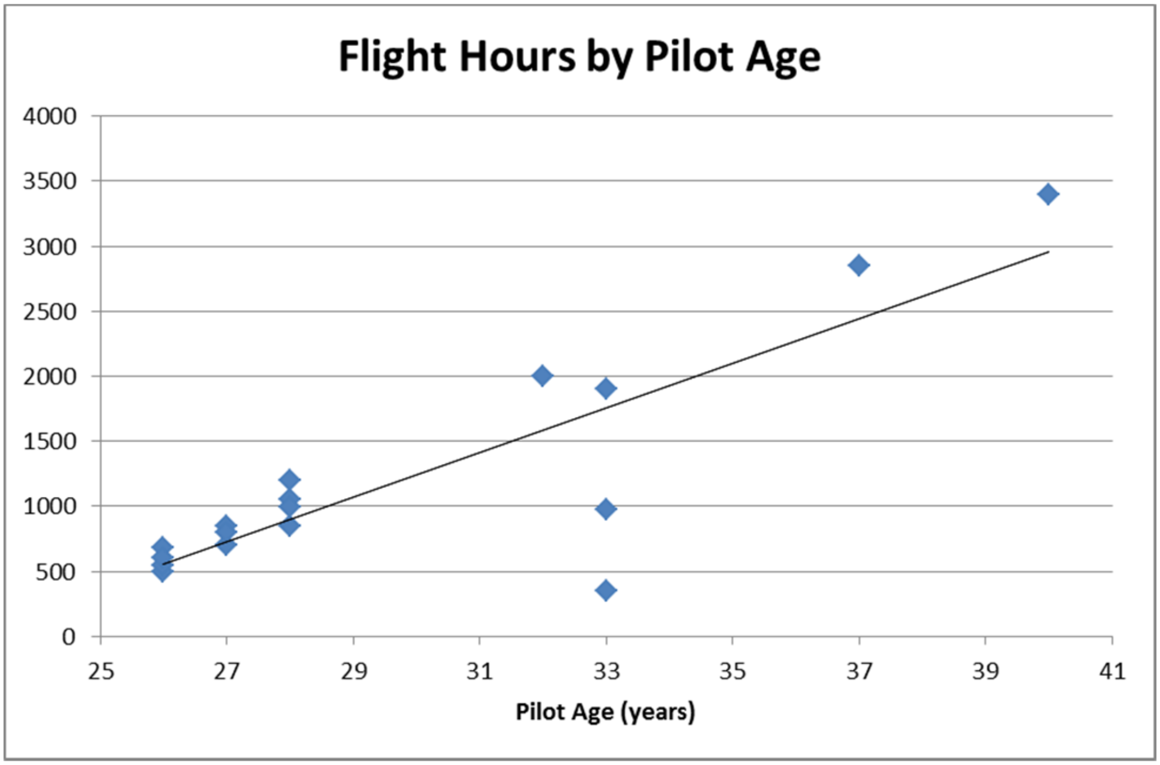


Figure 29. Total Flight Hours by Pilot Age

With age, processing speed (how fast the brain works) slows down, and older pilots are no exception from this pattern (Taylor Kennedy, Noda and Yesavage, (2007); Kennedy, Taylor, Reade, and Yesavage (2010).). Therefore, do the pilots in this study have a slower scan rate due to age? Or was the slower scan and longer fixation time due to greater experience? Table 23 outlines the Spearman's correlations between these eye

scan parameters and age (see Figures 26 and 27). Again, the correlations do not exist with any significance, and Figures 30 and 31 show a very weak linear relationship in both cases.

Trend	Spearman's RHO	P-value	Sig. Trend?	Type?
Median Dwell Times by Pilot Age	0.241	0.204	N	Positive
Scan Rates by Pilot Age	-0.139	0.317	N	Negative

Table 23. Association between Eye Tracking Parameters and Pilot Age

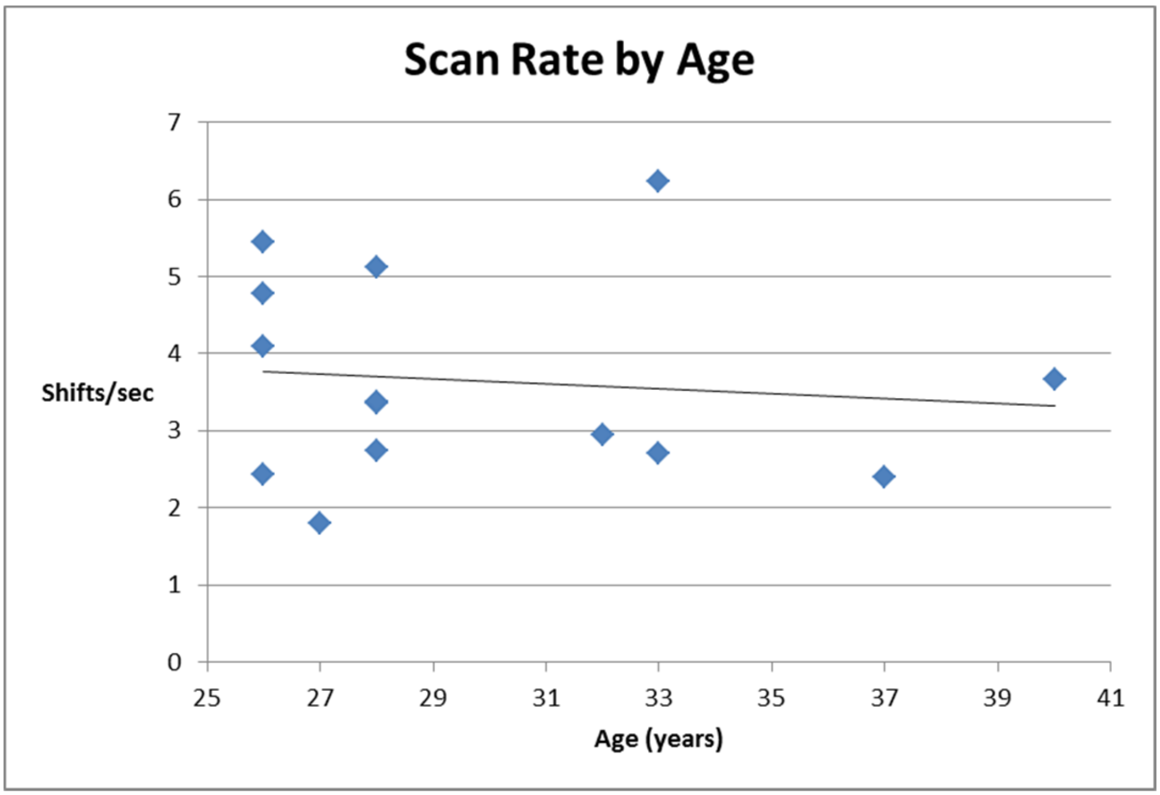


Figure 30. Scan Rate by Pilot Age

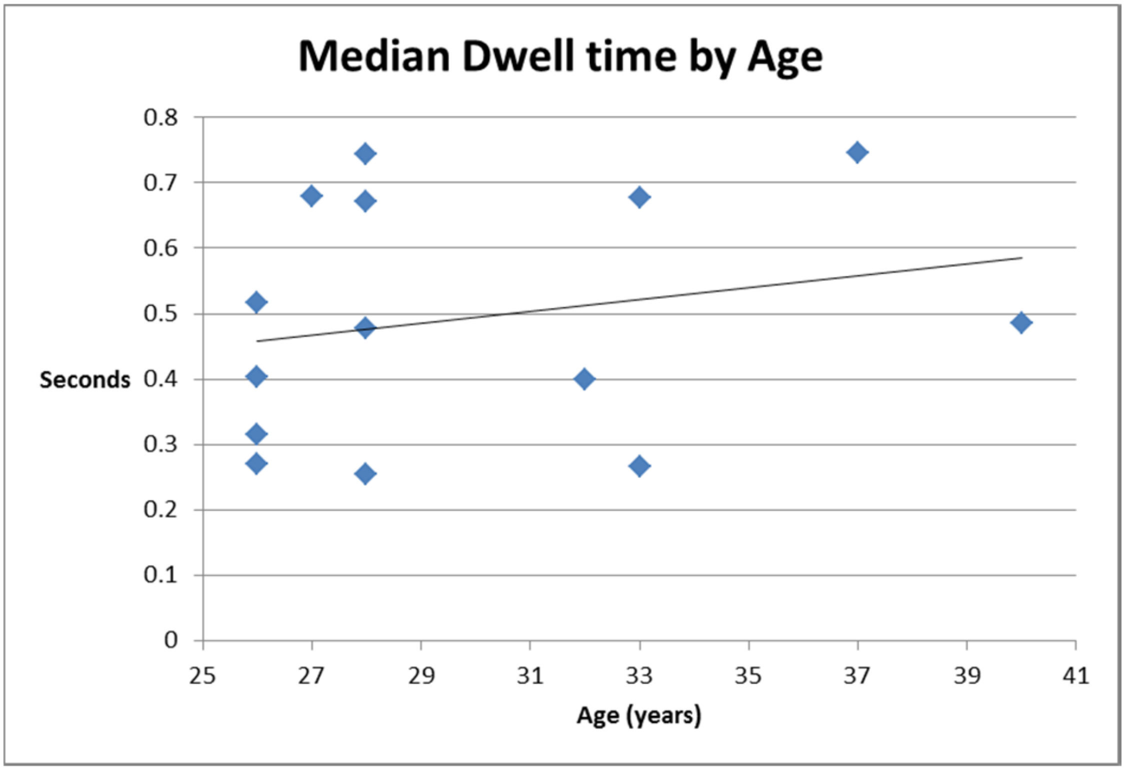


Figure 31. Median Time by Pilot Age

THIS PAGE LEFT INTENTIONALLY BLANK

IV. CONCLUSION AND DISCUSSION

A. DISCUSSION

The research focused on scanning patterns in flight regimes that are not optimal: flight at high speeds and low altitude levels. For military helicopter pilots, maintaining a low altitude is crucial to maintain survivability in hostile environments, reducing the amount of time an enemy has to detect, track, and possibly destroy the helicopter. These flight parameters are unique to helicopter aviation. It was our hope to gain an understanding of the unique scan characteristics that might present themselves in this challenging arena of flight by investigating the relationship between a helicopter pilot's experience level and his or her visual scan patterns during high-speed, low-level flight. A secondary goal of this thesis was to verify that FaceLab can be adapted for use in a simulator that is not in the laboratory environment.

B. FLIGHT PERFORMANCE RESULT

Flight performance typically is measured by adherence to flight path (i.e., RMS error in Caldwell, Jones, Carter and Caldwell, 1992), but Sullivan et al. (2011) found that levels of flight expertise did not predict RMS error of flight trajectory in helicopter overland navigation as well as it does in other aviation tasks. Based in part on the results in the Sullivan et al. paper, I decided to use mean altitude and altitude standard deviation as the main indicators of flight performance in this thesis, rather than RMS error. Pilots were instructed to stay between 100–300 feet above ground level and the average altitude and altitude standard deviation were the main measures of flight performance. Although all pilots maintained an average altitude of below 500 feet, more experienced pilots were better able to maintain a lower and more consistent altitude. These results are consistent with previous research findings that a pilot's experience expressed as total flight hours is positively associated with flight performance (Wickens et al., 2003, and Soliday and Schohan, 1965).

C. VISUAL SCAN RESULTS

Analyses of pilots' visual scans revealed a clear eye-scan pattern. Pilots spent most of their time scanning between out of the window and the instrument display. The map display, which was the aircraft diagnostics page, was not scanned often. The pilots were told not to expect adverse conditions with the aircraft, so the few scan shifts to and from the map display, combined with the relatively few fixations on the map display (Table 16), indicate that the pilots were not expecting anything to go wrong. Thus, with the exception of one pilot, they spent very little time worrying about the condition of the aircraft and concentrated on the navigation and maintaining the assigned flight parameters.

D. HYPOTHESES RESULTS

Results from tests of Hypotheses 1 and 2 were generally in the predicted directions, but did not reach significance. For Hypothesis 1, regarding the association between eye scan parameters and total flight time, the strongest trends were the amount of time pilots spend looking out of the window and looking at the instrument display. As flight experience increased, pilots spent more time looking at the instrument display than out of the window, swapping OTW scan time for more instrument scan time. Further evidence of this trend is that the more experienced pilots had more fixations on the instrument display than the less experienced pilots. These results agree with those found by Ottati et al. (1999) who found that novice pilots were more likely to fly scanning out of the window rather than relying on instrumentation to guide them through a navigational route. Sullivan et al. (2001) also found that the more experienced pilots scanned OTW less frequently than the less experienced pilots.

To explore this idea further, I distinguished between fixations that lasted longer than 70 milliseconds and those that lasted less than this time cutoff (labeled “no fixations”). This time cutoff was used to determine when a person was clearly fixating on something versus just a visual skim. A fixation is an event in which a pilot takes more time to gather information from the area of interest. A fixation is not “staring”; it is merely an event in which the pilot dwells on the area of interest in an attempt to interpret

the information from it. This event can happen during OTW, IP or MAP dwells—the pilot could have seen something of interest that might take more time to interpret, or the pilot chooses to be more deliberate in the time he scans each area of interest. Exploratory analyses revealed that the amount of “no fixation” events significantly decrease with regard to pilot experience. This trend is consistent with the result that the more experienced pilots had more fixations than the less experienced pilots. Both Bellenkes et al. (1997) and Karsarkis et al. (2001) found results that were similar; they found that experts tended to have more fixations.

Hypothesis 2 addressed the association between eye scan parameters and flight performance. A negative correlation between the average dwell percentages for OTW with altitude deviation showed that pilots who spent more time looking OTW had overall poorer performance in maintaining a constant altitude, a result similar to that found by Karsarkis (2001). No other correlations approached significance. The lack of strong correlations between flight performance and other eye scan parameters parallels the findings of Sullivan et al. (2011) who reported that eye scan parameters did not predict RMS error.

Hypothesis 3 was designed to investigate CFIT occurrence. However, no CFIT events occurred, so I was unable to test this hypothesis. With more subjects, the probability of a crash occurring may increase but, given the experience levels of the participants and the relatively low task load placed on them during the flight, such a crash was unlikely. These pilots were very familiar with the route because they had flown it many times on their routine training flights. Future research should consider making the simulated flight route sufficiently challenging to "induce" CFIT events.

E. EXPLORATORY ANALYSES

The relationship between scan patterns, average dwell time and pilot age was investigated in the exploratory analyses section of Chapter Three. As expected, there was a trend for a positive correlation between flight hours and pilot age. Older pilots typically have had more time to accrue flight hours (Taylor et al., 2007). Analyses revealed that the correlations between (1) the eye tracking parameters and total flight hours and (2) eye

tracking parameters and pilot age were very similar. The exceptions were two "older" pilots, a 33-year old with 350 hours and another 33-year old with 975 hours. These pilots may have entered flight school at an older age when compared to their peers, thereby falling a little behind their peers in terms of flight hours. The pilot with 350 hours had a scan rate and dwell time within a standard deviation of the group means, but the pilot with 975 hours had a scan rate and mean dwell time greater than one standard deviation difference above the group means.

Two case studies were investigated. Subject 18 had a higher than average altitude standard deviation than the other pilots in the study. The most experienced pilot in the study, Subject 8, had a large amount of dwell time on the Map display—a value far greater than any other pilot in the study. In both cases the data from FaceLab was cross-referenced with simulator performance data to investigate possible causes for differences in performance. While the results were inconclusive in both cases, two sources of data were successfully integrated and cross-referenced. The integration of the two systems was one of the goals of the study.

F. LIMITATIONS

One likely reason for the failure to reject the null hypotheses is the small sample size. Although 17 subjects participated in the study, the effective sample size ranged between 12 and 15 pilots, depending on the analyses. Two of the subjects' data were lost when the simulator's de-brief system failed to record their performance parameters. The FaceLab data for three more subjects was determined to be inaccurate, in that the eye tracking system was unable to record the pilot's eye and head movement for a significant duration of the flight. When the trend analyses were performed on the performance data or the eye tracking data separately, only 15 of the subjects' data were available for the performance evaluation, and only 14 of the subjects' data were available for trend analysis on the eye tracking data. When the eye tracking data was compared to the performance data, only 12 of the subject's data were available for testing.

G. STRENGTHS

The major strengths of this study were the demographic characteristics of the sample of pilots and the flight simulator used. The pilots represented a wide range of flight experience in terms of flight hours (min 350, max 3400) and were all in the midst of an operational flying tour. Thirteen of the pilots had all flown within a month of the trials, and only three pilots had more than one month away from flying (max two months since last flight). Many of the studies cited in Chapter One drew from populations consisting of only civilian pilots. This study was successful in acquiring pilots from a military helicopter community that specializes in low-level, high-speed flight.

TOFT-2, the simulator used in this study was actually being used for military helicopter training at the time this study was conducted. The simulator was also an exact replica of the aircraft flown by all of the pilots in this study. At the time of this study, TOFT-2 was still being used for training and evaluation of fleet pilots. All of the pilots in the study had experience in the kind of simulator TOFT-2 represented. Also, TOFT-2's video graphics were representative of the area where the pilots operate and train in: eastern San Diego. The research team was successful in showing that FaceLab could be installed in an operational fleet simulator and produce usable data for analysis. This was a major goal of the research effort.

These characteristics, the sample and the flight simulator, provide insight into how the use of psycho-physiological measures, such as eye tracking, can be used to aid training effectiveness specific to military helicopter missions, particularly in high-speed, low-level flight in visual conditions. The combination of flight simulator performance measures and eye-tracking parameters can be particularly useful in understanding why pilots exhibit certain performance levels. The studies by Bellenkes et al. (1997) and Ottati et al. (2001) both used military flight simulators in the same fashion, but only the study by Sanders et al. (1979) used a military simulator that was designed to emulate a helicopter (the UH-1). This study is one of the few that focuses only on military helicopter pilots flying in conditions that only they will encounter in their day-to-day operations.

H. IMPLICATIONS OF THE RESULTS

The information gained from understanding eye scan pattern during flight at high speeds and low altitude levels could be used in the development of a viable Heads-Up Display (HUD) for the MH-60S. Table 16 shows that pilots spent some of time (35%) fixating on the instrument display. Those who spent more time looking OTW were less likely to maintain constant altitude. Given that the instrument display provides valuable information, yet the pilot also needs to regularly look OTW, the HUD would greatly narrow the distance that pilots would have to scan between OTW and the instrument display. It would allow pilots to keep their scan outside while still gaining valuable flight and navigation data from the HUD. The instrument display information coupled with limited aircraft diagnostics data would greatly reduce the amount of time a pilot would have to divert his scan from the outside world. Indeed, Mumaw et al. (2001) came to similar conclusions and results from their study using informed instrumentation to set up training programs.

I. SUMMARY

In sum, results from this research may aid training effectiveness: spending more time looking at the instrument display and less time looking OTW aids maintenance of altitude in today's helicopter. Long looks OTW may be particularly detrimental in maintaining low-level altitude. Now guidelines can be created based on knowing how the more experienced pilots scan while flying at low altitude levels. Because it is apparent that navy helicopter pilots are flying approximately half of their total flight time over the land, it is critical to understand the scan patterns of the more experienced pilots while they are flying in this regime and pass that knowledge on, in the form of structured training, to the pilots of the future. Training syllabi are often “written in blood.” This is a common saying in the aviation world, describing how most training manuals are written by examining the mistakes of other aviators. Being able to write training manuals based on science and research is far preferable than waiting for another mishap to happen.

J. RECOMMENDATIONS

Several recommendations are made to improve any future studies in this area.

1) The amount of data collected throughout the duration of this study from the combined sources of the surveys, the simulator, and the FaceLab system presents more than enough material for further testing and comparison. This study focused primarily on experience level as defined by total flight hours. Additional analyses can be done comparing eye scan patterns with the amount of overland flying time, or the type of missions the pilots flew for the majority of their careers.

2) As discussed above, the small sample size was the main limitation of the study. There were many reasons the sample size was limited to 17 overall. Training commitments in the simulator, operational commitments of the squadron, and the fact that TOFT-2 was scheduled for decommissioning all contributed to the short amount of time allotted for the study. A dedicated time for the simulator use, in that the simulator would be used only to facilitate the study, is key to establishing a larger participant base. All commands involved were very supportive in lending pilots to the study. That being said, it was impossible for anyone to foresee the operational and training commitments that were placed on the pilots on a day-to-day basis. Some pilots were unable to make their scheduled times due to extended flights, or shifts in the flight schedule. A challenge for the next study team will be trying to eliminate these cancellations.

3) The simulated conditions used in this study were the same for every participant. All flights were conducted during the daytime hours, with no added weather effects. Further research could be conducted as to how pilots will scan given they are flying in inclement weather, at night, or a combination of inclement weather at night. This variation might give further insight into the efficiency or inefficiency of the layout of cockpit displays, given the changes seen when pilots are asked to fly in situations that increase the demands placed on them by the environment.

4) A goal should be to get eye-tracking technology integrated into an actual aircraft cockpit. This integration will require approval and coordination between various agencies within the Navy, but the benefits of having the system in an actual aircraft would be noteworthy. While modern simulators are realistic in their motion and visual graphics, there is no substitute for the real thing. The primary difference is that mistakes that would usually prove fatal in the aircraft are only result in the “red screen of death” in

the simulator. Pilots may fly the simulator differently than they would the actual aircraft. Any results obtained in any cockpit other than that of the actual aircraft will still have the term “simulated” in front of them, which is why this experiment is a crucial first step towards the incorporation of this technology into an actual aircraft, to see how helicopter pilots scan during actual flight.

5) Looking in to the future, with the hope that a Heads-Up-Display will be incorporated into rotary-winged aircraft, studies using FaceLab-type technology can be used to determine the benefits of using HUD technology. These results, when compared to the scan patterns analyzed from a non-HUD equipped aircraft, could add further evidence that HUDs will indeed improve flight safety by reducing pilot cognitive task loading.

APPENDIX A

A. QUESTIONNAIRES

Analysis of Helicopter Pilot Scan Techniques:
Simulation Exercise Evaluation
Flight Experience Questionnaire

Please provide the following information.

1. Profile Information

Age Gender Rank

2. The following questions ask about your flight experiences.

   Total flight hours:

   Overland hours:

   Branch of Service:

   Community:

   Years of aviation experience:

3. How many months has it been since your last flight?

4. How many months has it been since your last overland navigation flight?

5. Would you qualify the bulk of your flying experience as: (circle one)
   Maritime Overland

6. Describe your operational flying experience:

Analysis of Helicopter Pilot Scan Techniques:
Simulation Exercise Evaluation
Navigation Questionnaire

We are interested in learning about your navigation experience. The following questions ask about your navigation experiences.

1. To what extent have you participated in activities other than overland navigation that may contribute to improved navigation skills? (Examples may include sport orienteering, land navigation exercises, boy/girl scouts etc.)?

○	○	○	○	○
No Related Experience	Very Limited Related Experience	Limited Related Experience	Somewhat Significant Experience	Significant Related Experience

2. At your peak of currency, how would you rate your navigation skills in a low-level (below 200’ AGL) overland environment?

○	○	○	○	○
Poor	Fair	Average	Good	Excellent

3. If tasked today, how would you rate your navigation skills in a low-level (below 200’ AGL) overland environment?

○	○	○	○	○
Poor	Fair	Average	Good	Excellent

4. How much experience do you have with low-level navigation in mountainous desert terrain?

○	○	○	○	○
None	Very Little	Somewhat	Considerable	Extensive

5. How much low level navigation experience do you have in the Southern California operating area?

○ ○ ○ ○ ○

An Analysis of Helicopter Pilot Scan Techniques:
Simulation Exercise Evaluation
Post-Task Questionnaire

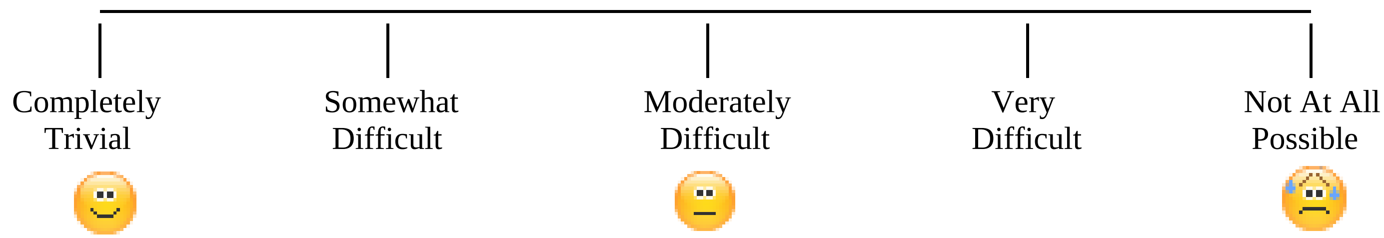
Please answer the questions below regarding how difficult you found the navigation and target detection tasks.

1. How difficult was it to navigate the route while maintaining the assigned parameters?

○	○	○	○	○
Not At All Difficult	Somewhat Difficult	Moderately Difficult	Very Difficult	Extremely Difficult

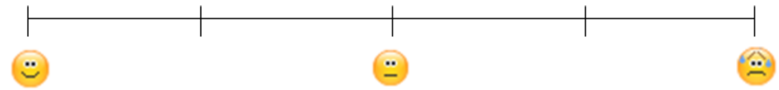
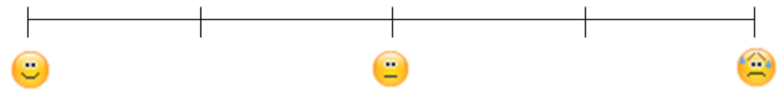
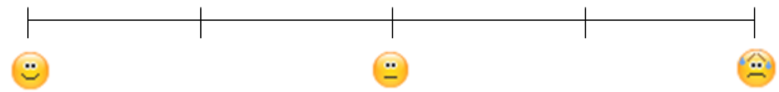
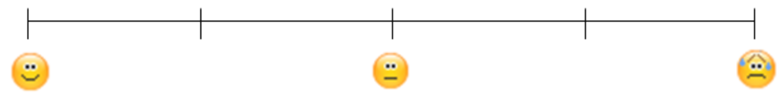
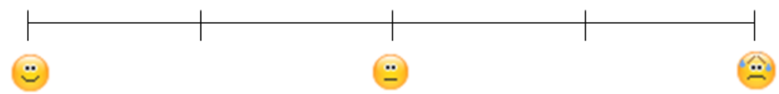
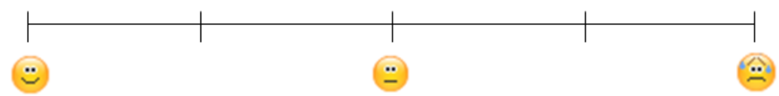
2. Describe any strategies that you used to stay on course and within the assigned flight parameters.

Please use the scale below to answer the questions 3–5.

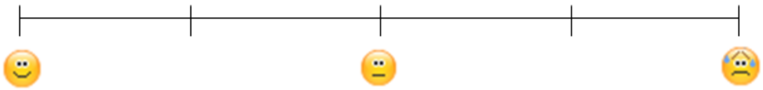


3. For each navigation leg on the route, please rate how difficult it was to navigate by referencing terrain. Place an “X” on the line that best describes your experience. No response is necessary for the shaded regions.

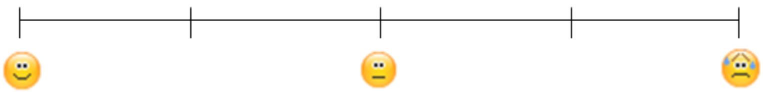
	Navigation Only	
Leg 1		
Leg 2		
Leg 3		
Leg 4		
Leg 5		

Leg 6		
Leg 7		
Leg 8		
Leg 9		
Leg 10		
Leg 11		
Leg 12		

4. How confident are you that you flew within the assigned parameters?

	Navigation	
	 Very Confident · Moderately Confident · Not At All Confident	

5. How confident are you that you correctly navigated the course?

	Navigation	
	 Very Confident · Moderately Confident · Not At All Confident	

APPENDIX B

A. [R] CODE FOR DATA ANALYSIS

[R] code used to process the Face Lab timing files.

```
# Function to process the timing files, function process.timingfiles:

function(filename){
# read in the file
thisline =
paste("C:/Users/Nex/Documents/NPS/thesis/Resultsandwriteups/inputfiles/",filename,sep
="")
thisinfile = read.csv(thisline, header = T)
attach(thisinfile)
# create the outfile for this subject
thisoutfile = paste("out_",filename,sep="")
thisoutlocation =
paste("C:/Users/Nex/Documents/NPS/thesis/Resultsandwriteups/outputfiles/",thisoutfile,
sep="")
# start creating the file for each category
#x = summary(FRAME_NUM)
#n = length(x)
#write(names(x),thisoutlocation,ncolumns = n,append=T,sep=",")
#write(x,thisoutlocation,ncolumns = n,append=T,sep=",")
colsinfile = length(names(thisinfile))
write(names(thisinfile),thisoutlocation, ncolumns = length(names(thisinfile)), append =
T, sep = ",")
n = length(FRAME_NUM)
for (i in 1:n){
	if (ANNOTATION_ID[i] != -1){
		thislinetowrite =
c(FRAME_NUM[i],EXPERIMENT_TIME[i],GMT_S[i],GMT_MS[i],DELAY[i],ANN
OTATION_ID[i])
		write(thislinetowrite,thisoutlocation,ncolumns = colsinfile,append = T, sep
= ",")
	}
}
}
#end of function
```

```
timinginfiles <-
c("Timing_subj7.csv","Timing_subj8.csv","Timing_subj11.csv","Timing_subj12a.csv","
Timing_subj12b.csv","Timing_subj12c.csv",
"Timing_subj13a.csv","Timing_subj13b.csv","Timing_subj14.csv","Timing_subj15.csv",
"Timing_subj16.csv","Timing_subj18a.csv","Timing_subj18b.csv",
"Timing_subj19.csv","Timing_subj20.csv","Timing_subj21a.csv","Timing_subj21b.csv",
"Timing_subj23.csv","Timing_subj25.csv","Timing_subj26.csv",
"Timing_subj28.csv")
timinginfiles <- c("Timing_subj22.csv")
m = length(timinginfiles)
for (i in 1:m){
        process.timingfiles(timinginfiles[i])
    cat("***!!!file ",timinginfiles[i]," complete!!!***","\n")
        }
```

[R] code used to process the Face Lab world view files.

```
# Function process.worldviewfiles: function for reading in, truncating, and producing a
summary of a world view

function(filename){
# read in the file
thisline =
paste("C:/Users/Nex/Documents/NPS/thesis/Resultsandwriteups/inputfiles/",filename,sep
="")
thisinfile = read.csv(thisline, header = T)
attach(thisinfile)
# create the outfile for this subject
thisoutfile = paste("out_",filename,sep="")
thisoutlocation =
paste("C:/Users/Nex/Documents/NPS/thesis/Resultsandwriteups/outputfiles/",thisoutfile,
sep="")
# start creating the file for each category
x = summary(ITEM_NAME)
n = length(x)
write(names(x),thisoutlocation,ncolumns = n,append=T,sep=",")
write(x,thisoutlocation,ncolumns = n,append=T,sep=",")
#x = summary(FRAME_NUM)
#n = length(x)
#write(names(x),thisoutlocation,ncolumns = n,append=T,sep=",")
#write(x,thisoutlocation,ncolumns = n,append=T,sep=",")
```

```
headers = c("Frame number","Change from", "To")
write(headers,thisoutlocation, ncolumns = 3, append = T, sep = ",")
n = length(FRAME_NUM)
thisitem = ITEM_NAME[1]
for (i in 2:n){
	if (ITEM_NAME[i] != thisitem){
		thislinetowrite = c(FRAME_NUM[i-1],thisitem,ITEM_NAME[i])
		write(thislinetowrite,thisoutlocation,ncolumns = 3,append = T, sep = ",")
	}
	thisitem = ITEM_NAME[i]
}
}
#end of function

# next, write a loop for all of the files using the worldview function to read them in:

#wvinfiles <- c("Wv-subj7.csv","Wv-subj8.csv","Wv-subj11.csv","Wv-subj12b.csv",
#"Wv-subj13a.csv","Wv-subj14.csv","Wv-subj15.csv","Wv-subj16.csv","Wv-
subj18b.csv",
wvinfiles <- c("Wv-subj19.csv","Wv-subj20.csv","Wv-subj21a.csv","Wv-
subj21b.csv","Wv-subj23.csv","Wv-subj25.csv","Wv-subj26.csv",
"Wv-subj28.csv","Wv-subj22.csv")

#wvinfiles <- c("Wv-subj11.csv")
m = length(wvinfiles)
for (i in 1:m){
	process.worldviewfiles(wvinfiles[i])
	cat("***!!!file ",wvinfiles[i]," complete!!!***","\n")
	}
```

THIS PAGE LEFT INTENTIONALLY BLANK

APPENDIX C

A. EQUATIONS

Trend analyses were done using the measures of rank correlation outlined in Conover's *Practical Nonparametric Statistics* (1999). The type of test for trends used was Spearman's *Rho* (Conover, 314). The tests used the data based on bi-variate samples to see if a trend existed as the one sample is ranked according to the other. *Rho* (ρ) is computed using the following equation (Conover, 315):

$$\rho = \frac{1-6\sum_{i=1}^{n}[R(X_i)-R(Y_i)]^2}{n(n^2-1)}$$

$R(X_i)$ represented the rank of the value from the sample X (the same held for $R(Y_i)$). The value *n* represented the sample size. A negative *Rho* value indicated an inverse relationship, a positive value indicates a direct relationship. For example, a negative *Rho* value indicated that altitude standard deviation decreased as pilot experience increased.

Helicopter group analyses utilized a two sample t-test assuming unequal variances, using the following equation (Devore, 337):

$$t = \frac{\bar{x}-\bar{y}-\Delta_0}{\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}}$$

$\bar{x}$ and $\bar{y}$ represented the means of each sample, and Δ_0 represented a value of the difference of the means of the two samples (in all cases for these tests, Δ_0 was 0). The values s_1 and s_2 represented the standard deviations of the two chosen sample groups. *m* and *n* were the sizes of each group.

The scan direction data was analyzed using paired t-tests in [R], which uses the following equation (Devore, 345):

$$t = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

$\bar{d}$ was the mean of the differences of the two samples. Δ_0 represented the value of the difference in the means to be tested (0). The standard deviation of the differences of the means of the two samples was represented by s_d in this equation. Finally, n represented the sample size, which was equal between the two groups.

LIST OF REFERENCES

Adams, C. (2010). Rotorcraft enhanced and synthetic vision. *Rotor and Wing, September 2010*, 32–37.

Baloh, R., Sills, A., Kumley, W., & Honrubia, V. (1975). Quantitative measurement of saccade amplitude, duration, and velocity. *Neurology, 25*, 1065–1070.

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: Their role of attentional flexibility and mental model development. *Aviation*, *Space, and Environmental Medicine, 68*(7), 569–579.

Bitton, D. (2008). Helicopter flight: The basics of preventing inadvertent IMC and CFIT. *Rotor, Winter 2008–2009*, 34–39.

Brehmer, B., & Dorner, D. (1993). Experiments with computer simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of field study. *Computers in Human Behavior, 9*, 171–184.

Burnside, B. L., & Throne, M. H. (2004). *Capabilities of future training support packages.* No. 1828. Fort Knox, KY: U.S. Army Research Institute for the Behavioral and Social Sciences.

Caldwell, J., Jones, H., Carter, D., & Caldwell, L. (1992). *The relationship between computer scoring and safety-pilot grading of flight performance.* Fort Rucker, AL: Biomedical Applications Research Division of the United States Army Aeromedical Research Laboratory.

Craig, Paul. (2001). *The killing zone* (pp. 1, 6, 304–320). New York: McGraw-Hill.

Conover, W.J. (1999). *Practical nonparametric statistics* (pp. 316). New York: John Wiley and Sons, Inc.

Crognale, M. (2008). Project summary of the final research report submitted to the federal aviation administration office of the chief scientist and technical advisor for human factors – general human factors program manager. Federal Aviation Administration, January 2008. 4–23. http://www.tc.faa.gov/logistics/grants/pdf/2005/05-G-018.pdf

Devore, J.L. (7th ed.), *Probability and statistics for engineering and the sciences*. San Luis Obispo, CA Brooks/Cole, 2009. 345–347.

Dorr, L., & Duquette, A. (2010). Fact sheet—helicopter emergency medical service safety. *Federal Aviation Administration News/Facts,* August 2011.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). New York: Wiley and Sons.

Duqette, A., & Dorr, L. (2010). Fact sheet—helicopter emergency medical safety. The Federal Aviation Administration. www.faa.gov/news/fact_sheets/news_story.cfm?newsId=6763

Evans, D. (2008). Unsaved—the deadly medical helicopter accident record. *Rotor, Air Safety, Winter 2008–2009,* 35–39.

Goodman, A., Folye, D. C., Hooey, B.L. & Wilson, J. R. (2003). Characterizing visual performance during approach and landing with and without synthetic vision display: A part task study. *Proceedings of the 2003 Conference on Human Modeling of Approach and Landing with Augmented Displays,* (NASA/CP-2003-212267), 71–89. Moffet Field, CA: NASA.

Hardiess, G., Gillner, S., & Mallot, H. A. (2008). Head and eye movements and the role of memory limitations in a visual search paradigm. *Journal of Vision, 8*(1)(7), 1–13. doi:10.1167/8.1.7

Harris, J. (1998) Sikorsky S-76 accident rates show overall decrease. Flight Safety Foundation, *Helicopter Safety,* 24:5, September-October 1998, 1–6

Harris, R., & Glover, B. (November, 1985). Effects of digital altimetry on pilot workload. (NASA Technical Memorandum 86424, TM-86424-1986-0004423). Savoy, Illinois: University of Illinois at Urbana-Champaign, Aviation Research Lab.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Science, 9*(4), 188–194.

Helleberg, J., Wickens, C., & Goh, J. (2003). Effects of data-link modality and display redundancy on pilot performance: An attention perspective. *The International Journal of Aviation Psychology, 13*(3), 189–210.

Helleberg, J., Wickens, C., & Goh, J. (2003). *Traffic and data link displays: Auditory? visual? or redundant? A visual scanning analysis.* Presented at the 12th International Symposium on Aviation Psychology. Dayton, OH.

Karacan, H., Cagiltay, K., & Tekman, H. G. (2010). Change detection in desktop virtual environments: An eye-tracking study. *Computers in human behavior, 26,* 1305–1313. doi:10.1016/j.chb.2010.04.002

Kasarskis, P., Stehwien J., Hickox, J., Aretz, A., & Wickens, C. (2001). *Comparison of expert and novice scan behaviors during VFR flight.* Presented at the 11th International Symposium on Aviation Psychology. Dayton, OH.

Kennedy, Q., Taylor, J. L., Reade, G., & Yesavage, J. (2010). Age and expertise effects in aviation decision making and flight control in a flight simulator. *Aviation, Space, and Environmental Medicine, 81,* 489–497.

Learmount, D. (2005). Honeywell's helicopter CFIT study dispels accident myths. *Flight International, 168,* 36.

Lohrenz, M. C., & Beck, M. R. (2010). Evidence of clutter avoidance in complex scenes. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting,* 1355–1359.

Mackworth, N. H. (1961). Researches on the measurement of human performance. In H. W. Sinaiko (Ed.), *Selected papers on human factors in the design & use of control systems* (pp. 174–331). New York: Dover Publications.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Maltz, M., & Shinar, D. (1999). Eye movements of younger and older drivers. *Human Factors: The Journal of Human Factors and Ergonomics Society, 41*(1), 15–25.

Marius't Hart, B., Vockeroth, J., Schumann, F., Bartl, K., Schneider, E., Konig, P., & Einhauser, W. (2009). Gaze allocation in natural stimuli: Comparing free exploration to head-fixed viewing conditions. *Visual Cognition, 17*(6), 1132–1158. doi:10.1080/13506280902812304

Marshall, S. (2007). Identifying cognitive state from eye metrics. *Aviation, Space and Environmental Medicine, 78, supplement 1,* 165–175.

Morelli, F., & Burton, P. A. (2009). The impact of induced stress upon selective attention in multiple object tracking. *Military Psychology*, *21*(1), 81–97. doi:10.1080/08995600802565769

Mumaw, R., Sarter, N., & Wickens, C. (2001). *Analysis of pilot's monitoring and performance on an automated flight deck.* Presented at the 11th International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University.

Nystrom, M. ,& Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking Data. *Behavioral Research Methods, February, 42*(1), 188–204.

Oman, C. M., Shebilske, W. L., Richards, J. T., Tubre, T. C., Beall, A. C., & Natapoff, A. (2002). Three dimensional spatial memory and learning in real and virtual environments. *Spatial Cognition and Computation, 2*(4), 355–372.

Ottati, L., Hickox, J., & Richter, J. (1999). Eye scan patterns of experienced and novice pilots during visual flight rules navigation. *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*, 1999.

Palinko, O., Kun, A., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. *Proceedings of the 2010 Symposium on Eye Tracking Research Applications, ETRA 10.* ACM Press, 2010. 141–144.

Pausch, R., Shackelford, A., & Proffitt, D. (1993). A user study comparing head-mounted and stationary displays. *IEEE Symposium on Research Frontiers in Virtual Reality,* 41–45.

Proctor, R. W., & Van Zandt, T. (2008). *Human factors in simple and complex systems* (2nd ed.). Boca Raton, FL: CRC Press: Taylor & Francis Group.

Sanders, M., Simmons, R., & Hofmann, M. (1979). Visual workload of the copilot/navigator during terrain flight. *Human Factors: The Journal of Human Factors and Ergonomics Society, 21,* 369–383.

Sarter, N., Mumaw, R., & Wickens, C. (2007). Pilots' monitoring strategies and performance on automated light decks: An empirical study combining behavioral and eye-tracking data. *Human Factors: The Journal of Human Factors and Ergonomics Society, 49*(3), 347–357.

Schriver, A., Morrow, D., Wickens, C., & Talleur, D. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors: The Journal of Human Factors and Ergonomics Society, 50*(6), 864–878.

Soliday, S., & Schohan, B. (1965). Task loading of pilots in simulated low-altitude high-speed flight. *Human Factors: The Journal of Human Factors and Ergonomics Society, 1965 7*(45), 45–53.

Sullivan, J., Yang, J., Day, M., & Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan pattern. *Aviation, Space, and Environmental Medicine, 82,* 871–878.

Svensson, E., Angleborg-Thanderez, M., Sjoberg, L., & Olsson, S. (1997) Information complexity-mental workload and performance in combat aircraft. *Ergonomics, 40,* 362–380.

Taylor, J. L., Kennedy, Q., Noda, A., & Yesavage. J. A. (2007). Pilot age and expertise predict flight simulator performance: A three-year longitudinal study. *Neurology, 68*(9), 648–654.

Teeter, P. (2011). *R Cookbook*. Sebastopol, CA: O'Reilly Media. Xiii.

Tiwari, T., Singh, A. L., & Singh, I. L. (July 2009). Task demand and workload: Effects on vigilance performance and stress. *Journal of the Indian Academy of Applied Psychology, 35*(2), 265–275.

Tole, J., Stephens, A., Harris, R., & Ephrath, A. (1982). Quantification of pilot workload via instrument scan. *Workshop for Flight Testing to Identify Pilot Workload Dynamics.* Edwards AFB, CA: Air Force Flight Test Center, Edwards Air Force Base.

Tvaryanas, A. (September 2003). Visual scan patterns during simulated control of uninhabited aerial vehicle (UAV). *Aviation, Space, and Environmental Medicine, 75*, 531–538.

Wickens, C. (2001). *Aviation safety and the psychology of surprise.* Proceedings of the 11th International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University.

Wickens, C., Dixon, S., Goh, J., & Hammer, B. (March 2005). Pilot dependence on imperfect diagnostic automation in simulated UAV flights: An attention visual scanning analysis. (Technical Report AHFD-05-02/MAAD-05-02). Savoy, IL: The University of Illinois at Urbana-Champaign, Aviation Human Factors Division, Institute of Aviation.

Wickens, C., Goh, J., Helleberg, J., & Talleur, D., (June 2002). Modality differences in advanced cockpit displays: Comparing auditory vision and redundancy for navigational communications and traffic awareness. (Technical Report ARL-02-8/NASA-02-6). Savoy, IL: University of Illinois at Urbana-Champaign, Aviation Research Lab.

Wickens, C., Goh, J., Helleberg, J., Horrey, W., & Talleur, D. (2003). Attention models of multitask performance using advanced display technology. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 45:*360. Doi: 10.1518/hfes.45.3.27250

Williges, B., Roscoe, S., & Williges, R. (1973). Synthetic flight training revisited. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 15,* 543–560.

Woodman, G. F., & Luck, S. (2010). Why is information displaced from visual working memory during visual search? *Visual Cognition, 18*(2), 275–295. doi:10.1080/13506280902734326

Yan L., Hsueh, P., Lai, J., Sangin, M., Nussli, M.,& Dillenbourg, P. (2008). *Who is the expert? Analyzing gaze data to predict expertise level in collaborative applications.* Presented at IEEE International Conference on Multimedia and Expo, June 2009. 898–901.

Yesavage, J., Otto Leirer, V., Denari, M., & Hollister, L. (November 1985). Carry-over effects of marijuana intoxication on aircraft pilot performance: Preliminary report. *American Journal of Psychiatry*, *142*(11), 1325–1329.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Commander, Helicopter Sea Combat Wing Pacific Fleet
   COMHELSEACOMBATWINGPAC
   Naval Air Station North Island
   San Diego, California

4. Commanding Officer, Helicopter Sea Combat Squadron Twenty-One
   HELSEACOMBATRON TWO ONE
   Naval Air Station North Island
   San Diego, California

5. Commanding Officer, Helicopter Sea Combat Squadron Three
   HELSEACOMBATRON THREE
   Naval Air Station North Island
   San Diego, California

6. Commanding Officer, Helicopter Anti-Submarine Squadron Four
   HELANTISUBRON FOUR
   HELSEACOMBATRON FOUR
   HS-4 UNIT 25172
   FPO AP 96601-5704